

Gopalakrishna, Goutham

**Working Paper**

## Robust test of long run risk and valuation risk model

Quaderni - Working Paper DSE, No. 1107

**Provided in Cooperation with:**

University of Bologna, Department of Economics

*Suggested Citation:* Gopalakrishna, Goutham (2017) : Robust test of long run risk and valuation risk model, Quaderni - Working Paper DSE, No. 1107, Alma Mater Studiorum - Università di Bologna, Dipartimento di Scienze Economiche (DSE), Bologna, <https://doi.org/10.6092/unibo/amsacta/5702>

This Version is available at:

<https://hdl.handle.net/10419/177761>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*



<https://creativecommons.org/licenses/by-nc/3.0/>



ISSN 2282-6483

Alma Mater Studiorum - Università di Bologna  
DEPARTMENT OF ECONOMICS

## **Robust test of Long Run Risk and Valuation risk model**

Goutham Gopalakrishna

*Quaderni - Working Paper DSE N°1107*



# Robust test of Long Run Risk and Valuation risk model

Goutham Gopalakrishna\*

September 25, 2017

## Abstract

This paper tests the long run risk and valuation risk model using a robust estimation procedure. The persistent long run component of consumption growth process is proxied by a news based index that is created using a random forest algorithm. This news index is shown to predict aggregate long term consumption growth with an R-square of 57% and is robust to inclusion of other commonly used predictors. I theoretically derive an estimatable bias term in adjusted Euler equation of the model that arises due to measurement error in consumption data and show that this bias term is non-zero. Using a three pass estimation procedure that accounts for this bias, I show that the long run risk and valuation risk model fails to explain cross section of equity returns. This contrasts to the results from regular two pass Fama-MacBeth estimation procedure that implies that the same model explains the cross section of asset returns with statistically significant risk premia estimates.

**Keywords:** Long run risk, Valuation risk, Machine Learning, Three pass filter, Media, Consumption

**JEL:** G12, E21, E44

---

\*University of Bologna, Department of Economics; Piazza Scaravilli 2, 40126 Bologna, Italy. E-mail: goutha.gopalakrishna2@unibo.it. I thank Prof. Laura Bottazzi for valuable guidance and support. I also thank Prof. Sergio Pastorello for inputs. All errors are my own.

# 1 Introduction

The inability of the asset pricing models to reconcile the empirical facts between the macroeconomic and financial data has been an open issue for more than a couple of decades (Cochrane 2016 [9], Hansen and Singleton 1982 [17]). The US equity market demand a premium of 6.9% over the risk free rate and the standard consumption based asset pricing models do not justify such a large premium. As Mehra (2003)[24] pointed out, the US stock market, with Standard and Poor average real annual yield as proxy, has delivered a return of 7.9% over the last 110 years which is considerably higher than the average annual real T-bill rate of 1% during the same time span. The crux of the equity premium puzzle lies in the inability to reconcile the stock return premium with the standard representative agent macroeconomic model with CRRA (Constant Relative Risk Aversion) preference. The obvious shortcoming of a CRRA framework is that the time preference and risk preference are coupled together. One of the most influential models among an arsenal of papers that aim to explain the cross section of asset returns is the Long Run Risk model by Bansal and Yaron (2004) [3]. The framework used in the Long Run Risk model is different from the other modes like Habit formation by Campbell and Cochrane (1999)[8], or ICAPM (Intertemporal Capital Asset Pricing Model) by Merton (1973)[25] in the sense that a small but highly persistent component in consumption and dividend growth rate takes central importance in model development. Bansal and Yaron (2004) used the Epstein Zin utility and assumed (i) persistent component in the consumption and dividend process, (ii) fluctuating consumption volatility. The key assumption of the long run risk framework is that a better long run prospect of the economy raises the equity prices, even though the consumption may remain unaffected. In the long run risk model, the uncertainty is loaded to the supply side of the economy.

A recent contribution in this area is the Valuation Risk model by Albuquerque et al (2015)[1] who use demand side shocks instead of the supply side shocks to model the asset prices. Since most of the classical models including long run risk models load all shocks into the supply side of economy, Albuquerque et al (2015)[1] take a different perspective and introduce time preference shocks in the Epstein-Zin utility framework. This model not only fits the equity data well, but also the term structure of bond returns better than the long run risk model. The underlying framework and assumption in these two models is pretty much the same - (i) a representative agent Epstein Zin utility, (ii) a small but highly persistent component in consumption growth process. While the long run risk has persistent long run component entering the consumption and dividend growth processes, the valuation risk model has a persistent time preference risk entering the consumption growth process. Following Schorfheide et al (2016)[33], I combine the long run risk and valuation risk model into an unifying framework such that the consumption growth process contains both the long run component as well as the time preference component. The preliminary evidence provided by Schorfheide et al (2016)[33] show a strong support for the model in terms of its ability to explain the cross section of asset returns. Similar result is obtained by Liu and

Matthies (2016)[23] who test only for the long run risk model. The long run component and time preference component are essentially latent variables which means that they are not observed. The approach taken by Schorfheide et al (2016)[33] to recover the latent components is a Bayesian mixed frequency approach while Liu and Mattheis (2016)[23] use a news based textual approach. The news based textual approach uses frequency count of number of news articles published each month that contains the keywords such as 'consumption' and 'gross domestic product'. However, it is a noisy measure since there are many news articles that contain these keywords but are not related to the state of the economy. While I adopt the news based methodology from Liu and Mattheis (2016)[23], I use machine learning tools to *clean* the index in the sense that the news based index is created using only the articles that are relevant to the state of economy with a reasonable approximation.

As far as the empirics are concerned, in order to test the model, one has to use consumption data to start with. The problem of measurement error in consumption is widely documented in prior literature. Some of the recent papers have used proxies such as electricity consumption (Da and Yun, 2010)[10], and data on garbage produced (Savov, 2011)[32]. The problem in using proxies is that one cannot concretely establish a mapping from consumption data to these proxies. Hence I take a different approach and theoretically solve for measurement error component in the estimation equations. Vissing-Jorgensen (1998)[34], and Brav et al (2002)[6] use a multiplicative measurement error model since this will retain the Euler equations obtained from solving the optimization problem. But using multiplicative error model leads to conclusion that only discount factor gets affected, and therefore is rather a restrictive assumption. I instead use an additive measurement error model which allows us to modify the testable Euler equation. Recently De (2015)[11] used an additive measurement error model in a CRRA framework and found that the measurement error component helps to solve the equity premium puzzle to an extent. To my knowledge, such additive measurement error model has not been implemented for Epstein-Zin utility in a Long run risk framework. Implementing it would provide us two advantages: (i) we can obtain a modified Euler equation which will take into account the measurement error problem, (ii) we can estimate the bias to have a sense of seriousness of neglecting the measurement error in macroeconomic data. The bias term that I derive is shown to be dependent on the utility parameters, risk free rate, and return of assets. I use utility parameter estimates obtained in Schorfheide et al (2016)[33] and show that the magnitude of bias is non-zero.

Since the modified Euler equation has a bias term, the canonical asset pricing equation  $1 = E[MR]$  does not hold anymore. I derive the cross sectional asset price testing equations in the presence of bias induced from measurement error and show that the bias term enters as asset specific heterogeneity term in the Fama MacBeth cross sectional regression equations. Therefore, ignoring the bias term and using OLS estimates to obtain the risk premia will lead to biased estimates. To correctly obtain the risk premia estimates, I use the three pass estimation technique proposed by

Giglio and Xiu (2017)[14] which not only takes into account the omitted bias variable, but also other omitted variables which could possibly exist.

The key results of the paper are as follows. The news index explains the aggregate consumption growth 7 years down the line with an adjusted R-squared of 57%. The index is robust to inclusion of other variables such as price dividend growth, realized industrial production variance, and labor income growth. The bias in the Euler equation of the model that arises due to measurement error is non-zero. Using the three pass estimation to account for this bias leads to a conclusion that the long run risk and valuation risk factors are spurious and do not explain the cross section of asset returns. This is in stark contrast to the conclusion arrived using the standard two pass estimation technique which implies that the long run risk and valuation risk factors are statistically significant in explaining the cross section of asset returns. Therefore, there is a lack of support for the long run risk and valuation risk model once we take measurement error into account and use robust risk premia estimation method.

Section 2 introduces the model set up<sup>1</sup>. This is followed by section 3 that explains the Data source. Section 4 presents the methodology and empirical results. Section 4.1 explains how the latent long run component in consumption growth process is recovered. The recovery of state variables are demonstrated in section 4.2. The estimates of bias in Euler equation is provided in section 4.3. At this point, a need for robust estimation procedure that takes into account measurement error becomes clear. Section 4.4 presents the robust estimation of risk premia using three pass approach and compares it against the estimates from standard two pass approach. Section 5 concludes the paper. The appendix A provides the mathematical derivation of the model and supporting figures and tables.

## 2 Model

A representative agent model with endowment economy and preferences characterized by Epstein and Zin (1989) is considered. The agent, holding recursive preferences, maximizes the lifetime utility by solving

$$U_t = \max_{C_t} \{ \lambda_t C_t^{1-\frac{1}{\psi}} + \delta ([U_{t+1}^*])^{1-\frac{1}{\psi}} \}^{\frac{1}{1-\frac{1}{\psi}}}$$

subject to the budget constraint,

$$W_{t+1} = (W_t - C_t)R_{c,t+1}$$

$C_t$  is the consumption stream,  $\psi$  is the Inter-temporal elasticity of substitution (IES), and  $\delta$  refers to the agent's time preference.  $R_{c,t+1}$  is the return on consumption claim,  $U_{t+1}^*$  is the

---

<sup>1</sup>I present the model and key equations in this section. Detailed derivation of solution to the model is provided in Appendix A

certainty equivalent value of lifetime utility from  $t + 1$  such that

$$(U_{t+1}^*)^{1-\gamma} = E_t(U_{t+1}^{1-\gamma})$$

Risk aversion parameter is given by  $\gamma$ . When  $\gamma = 1/\psi$ , the recursive preferences reduces to CRRA preferences. The demand shock is captured by the changes in time discount rate of investors. The variable  $\Lambda_{t+1} = \lambda_{t+1}/\lambda_t$  is the time preference shock that captures how the investors trade off current period utility versus future period utility. Let  $x_{\Lambda,t}$  denote the log time preference shock, which follows the law of motion

$$x_{\Lambda,t+1} = \rho_{\Lambda}x_{\Lambda,t} + \sigma_{\Lambda}\epsilon_{t+1}^{\Lambda}$$

where  $\epsilon_{t+1}^{\Lambda} \sim N(0, 1)$ . The law of motion for consumption process is a modified version of Bansal Yaron type since the innovation to the time preference shock also enters the consumption growth process. The state process is given by

$$\begin{aligned} g_{t+1} &= \mu_c + x_{c,t} + \pi_{c,\lambda}\epsilon_{t+1}^{\Lambda} + \sigma_{c,t}\eta_{t+1} \\ x_{c,t+1} &= \rho x_t + \varphi_e \sigma_{x,t} e_{t+1} \\ x_{\Lambda,t+1} &= \rho_{\Lambda}x_{\Lambda,t} + \sigma_{\Lambda}\epsilon_{t+1}^{\Lambda} \\ \sigma_{k,t+1}^2 &= \nu + \sigma_{k,t}^2 + \phi_0\zeta_t + \phi_1\zeta_{t-1} + \omega_{t+1}, \quad k \in \{c, x\} \\ \eta_t, e_t, \epsilon_t, \zeta_t &\sim \text{i.i.d standard-normal random variable} \end{aligned}$$

where  $g_{t+1}$  is the consumption growth,  $\pi_{c,\lambda}\epsilon_{t+1}^{\Lambda}$  is the scaled innovation to the time preference shock,  $x_{c,t}$  is the long run variable that follows an autoregressive process,  $x_{\Lambda,t}$  is the time preference shock that follows an autoregressive process,  $\sigma_{c,t}$ , and  $\sigma_{x,t}$  is the time varying component in the stochastic volatility of consumption growth  $g_{t+1}$ , and long run variable  $x_{c,t+1}$  respectively. A stochastic volatility is evidenced in the post war macroeconomic and financial data which makes it a reasonable assumption.  $\sigma_{c,t+1}^2$  and  $\sigma_{x,t+1}^2$  follows an AR(I)MA process of order (0,1,2), or in other words a simple exponential smoothing process<sup>2</sup>. The state variables in the model are  $\mathbf{S}_t = [x_{c,t} \quad x_t^{\Lambda} \quad \sigma_{x,t}^2 \quad \sigma_{c,t}^2]$

The stochastic discount factor given the Epstein Zin preferences with valuation risk embedded can be derived as <sup>3</sup>

$$m_{t+1} = \theta \ln \delta + \theta \ln \frac{\lambda_{t+1}}{\lambda_t} - \frac{\theta}{\psi} g_{t+1} + (\theta - 1)r_{a,t+1} \quad (1)$$

where  $\theta$  is given by  $\theta = \frac{1-\gamma}{1-\frac{1}{\psi}}$  with  $\gamma \geq 0$ . The stochastic discount factor  $m_{t+1} \equiv \ln M_{t+1}$  contains measurement error.

## Deriving Adjusted Euler Equation

I assume that observed consumption growth,  $\tilde{C}_{t+1}$  contains measurement error and  $\tilde{C}_{t+1} = C_{t+1} + \xi_{t+1}$ , where  $\xi_{t+1} = C_{t+1}^0 \epsilon_{t+1}$ . Similarly,  $\tilde{C}_t = C_t + \xi_t$ , where  $\xi_t = C_t^0 \epsilon_t$ . Also,  $\epsilon_{t+1}, \epsilon_t \sim$

<sup>2</sup>I consider AR(I)MA process in general and choose the best model according to AIC, and BIC value. I use *auto.arima* package in R that uses Hyndman-Khandakar algorithm for automatic ARIMA modeling

<sup>3</sup>see Appendix A for derivation

$i.i.dN(0, \sigma_\epsilon^2)$ .  $C_{t+1}$  and  $C_t$  are true consumption series which are not observed. This forms an additive consumption error model with heteroskedastic noise <sup>4</sup>. In the absence of measurement error, the Euler equation is given by  $1 = E_t[M_{t+1}R_{i,t+1}]$ . This pricing equation holds when the stochastic discount factor  $M_{t+1}$  is a function of true consumption data. When observed consumption data is noisy, we do not know what the true consumption data is and hence we require a modified Euler equation which is testable with the observed noisy consumption series. The Euler equation under measurement error is given as  $E_t[M_{t+1}R_{i,t+1}] = p$ . If  $p \neq 1$ , the measurement error affects the Euler equation.

To derive the modified testable Euler equation, I perform the following steps:

- Take the noisy consumption growth series  $\frac{\tilde{C}_{t+1}}{\tilde{C}_t}$  and expand it around the points  $E[\tilde{C}_{t+1}] = C_{t+1}^\circ$  and  $E[\tilde{C}_t] = C_t^\circ$  until the second order in Taylor Series expansion <sup>5</sup>
- Substitute the linearized expression for noisy consumption growth in the corresponding Euler equation  $E_t[M_{t+1}R_{i,t+1}] = p$ .
- Similarly, take the true (unobserved) consumption growth series  $(\frac{C_{t+1}}{C_t})$  and expand it around points  $E[C_{t+1}] = C_{t+1}^\circ$  and  $E[C_t] = C_t^\circ$  <sup>6</sup>.
- Substitute the linearized expression for true consumption growth in the corresponding Euler equation  $E_t[M_{t+1}R_{i,t+1}] = 1$ .
- Subtract the two Euler equations and obtain the bias term  $p$  as a function of utility parameters and returns. Substituting this  $p$  in the Euler equation with measurement error, we get the modified Euler equation <sup>7</sup>

$$\delta^\theta E_t \left[ \left( \frac{\lambda_{t+1}}{\lambda_t} \right)^\theta R_{i,t+1}^{\theta-1} R_{i,t+1} \left( \frac{\tilde{C}_{t+1}}{\tilde{C}_t} \right)^{-\frac{\theta}{\psi}} \right] \approx 1 + \delta^\theta \left( \frac{\theta}{\psi} \right)^2 \sigma_\epsilon^2 E_t \left[ \left( \frac{\lambda_{t+1}}{\lambda_t} \right)^\theta R_{a,t+1}^{\theta-1} R_{i,t+1} \right] \quad (2)$$

$$\delta^\theta E_t \left[ \left( \frac{\lambda_{t+1}}{\lambda_t} \right)^\theta R_{i,t+1}^{\theta-1} R_{i,t+1} \left( \frac{\tilde{C}_{t+1}}{\tilde{C}_t} \right)^{-\frac{\theta}{\psi}} \right] \approx 1 + \Pi \left[ \Omega, R_{i,t+1} \right] \quad (3)$$

The term  $\Pi \left[ \Omega, R_{i,t+1} \right]$  is the bias that is induced due to the measurement error <sup>8</sup>. The bias is a function of utility parameters (risk aversion  $\gamma$ , inter-temporal elasticity of substitution  $\psi$ ), discount rate ( $\delta$ ), time preference shock  $\frac{\lambda_{t+1}}{\lambda_t}$ , the aggregate wealth portfolio return ( $R_{a,t+1}$ ), the asset return ( $R_{i,t+1}$ ), and the variance of measurement error ( $\sigma_\epsilon^2$ ).

---

<sup>4</sup>This assumption is reasonable since consumption data has a trend

<sup>5</sup>Given the highly non-linear nature of the Euler equation, I expand up to two terms. I assume that the higher order terms have a negligible impact.

<sup>6</sup>I expand until the second order in Taylor Series expansion, again assuming that terms higher than second order have negligible impact

<sup>7</sup>derivation given in Appendix A

<sup>8</sup>It is assumed that the measurement error in observed consumption growth  $\tilde{C}_t$  is such that it is never zero in order to prevent the Euler equation from blowing up. In order to avoid this, one can use a truncated distribution but assuming the measurement errors are not so large that the consumption becomes exactly zero is a possibility that I omit.



## Solving the model

To characterize the stochastic discount factor we need to solve for return on asset  $r_{a,t+1}$ , which is endogenous in the model. I make use of the log-linearized return formula as in Campbell and Shiller (1988) and write the return of an asset in the form

$$r_{a,t+1} = \kappa_0 + \kappa_1 z_{t+1} - z_t + g_{t+1} \quad (4)$$

where  $z_t$  is defined as the log price consumption ratio ( $z_{t+1} = P_{c,t+1}/C_{t+1}$ ). The coefficients  $\kappa_0$  and  $\kappa_1$  are functions of unconditional expectations of price consumption ratio  $z_t$ . I use method of undetermined coefficients to solve for the return process. A conjecture for  $z_t$  is

$$z_t = A_0 + A_1 x_{c,t+1} + A_2 x_{t+1}^\Lambda + A_3 \sigma_{c,t+1}^2 + A_4 \sigma_{x,t+1}^2 \quad (5)$$

Similarly, the return on market portfolio and price dividend ratio is given by

$$r_{m,t+1} = \kappa_{0,m} + \kappa_{1,m} z_{m,t+1} - z_{m,t} + g_{d,t+1}$$

where  $g_{d,t+1}$  is the dividend yield growth. The price dividend ratio for market takes the affine form

$$z_{m,t} = A_{m,0} + A_{m,1} x_{c,t+1} + A_{m,2} x_{t+1}^\Lambda + A_{m,3} \sigma_{c,t+1}^2 + A_{m,4} \sigma_{x,t+1}^2$$

I first plug in equation (5) in equation (4) to write the return on asset in linear relation to underlying state variables. I then plug this resulting equation, as well as the SDF equation (1) in the Euler Equation (3). This leads us to writing the SDF in linear relation to the underlying state variables. The innovation to SDF is then given by,

$$m_t - E_t(m_{t+1}) = \theta [\sigma_\Lambda \epsilon_{t+1}^\Lambda] - \frac{\theta}{\psi} [\pi_{c,\lambda} \epsilon_{t+1}^\Lambda + \sigma_t \eta_{t+1}] + (\theta - 1)[r_{a,t+1} - E_t(r_{a,t+1})] \quad (6)$$

The risk free rate can be derived as <sup>9</sup>

$$\begin{aligned} r_{f,t} = & -\theta \ln \delta + \frac{\theta}{\psi} (\mu_C + x_{c,t}) - \theta \rho_\Lambda \ln \frac{\lambda_t}{\lambda_{t+1}} - (\theta - 1) \{ k + 0 + k_1 A_0 + k_1 A_1 \rho x_{c,t} + k_1 A_2 \rho x_t^\Lambda \\ & + k_1 A_3 E_t(\sigma_{c,t+1}^2) + k_1 A_4 E_t(\sigma_{x,t+1}^2) - A_0 - A_1 x_{c,t} - A_2 x_t^\Lambda - A_3 \sigma_{c,t}^2 - A_4 \sigma_{x,t}^2 + x_{c,t} + \mu_c \} \\ & - 0.5 \{ \theta^2 \text{Var} \sigma_\Lambda^2 + (\frac{\theta}{\psi})^2 (\pi_{c,\lambda}^2 + \sigma_{c,t}^2 + (\theta - 1)^2 \text{Var}(r_{a,t+1})) \} + \Pi[\Omega, R_{f,t+1}] \end{aligned}$$

The risk free rate is also linear in the state variables  $\mathbf{S}_t = [x_{c,t} \quad x_t^\Lambda \quad \sigma_{x,t}^2 \quad \sigma_{c,t}^2]$ . The derivation is explained in detail in Appendix A. The important thing is that the stochastic discount factor, and risk free rate can be represented linearly in terms of the risk factors. Given these expressions, the variance of the stochastic discount factor can be written in terms of the variance of the other risk factors in the model as

$$\begin{aligned} \text{Var}(m_{t+1}) = & \lambda_{m,\eta}^2 \text{Var}(g_{t+1}) + \lambda_{m,\lambda}^2 \text{Var}(x_{t+1}^\Lambda) + \lambda_{m,x}^2 \text{Var}(x_{c,t+1}) \\ & + \lambda_{m,\sigma_c}^2 \text{Var}(\sigma_{c,t+1}^2) + \lambda_{m,\sigma_x}^2 \text{Var}(\sigma_{x,t+1}^2) \end{aligned} \quad (7)$$

---

<sup>9</sup>see Appendix A for derivation.  $r_{f,t}$  depends linearly on State variables and  $\text{Var}(r_{a,t+1})$ . But  $\text{Var}(r_{a,t+1})$  is linear in state variables and therefore  $r_{f,t}$  is linear in State variables.

where the  $\lambda$ 's are given by,

$$\begin{aligned}\lambda_{m,\eta} &= (\theta - 1) - \frac{\theta}{\psi} \\ \lambda_{m,\lambda} &= \theta + k_1 A_2 (\theta - 1) \\ \lambda_{m,x} &= k_1 A_1 (\theta - 1) \\ \lambda_{m,\sigma_c} &= A_3 (\theta - 1) k_1 \\ \lambda_{m,\sigma_x} &= A_4 (\theta - 1) k_1\end{aligned}$$

Plugging these  $\lambda$ 's into the innovation in the stochastic discount factor (6), we get

$$\begin{aligned}m_t - E_t(m_{t+1}) &= [\sigma_\Lambda \lambda_{m,\lambda} + \pi_{c,\lambda} \lambda_{m,\eta}] \epsilon_{t+1}^\Lambda + \lambda_{m,\eta} \sigma_{c,t} \eta_{t+1} + \lambda_{m,x} \varphi_e \sigma_{x,t} e_{t+1} \\ &\quad + \lambda_{m,\sigma_c} \omega_{c,t+1} + \lambda_{m,\sigma_x} \omega_{x,t+1}\end{aligned}$$

We notice that the proportion of  $\epsilon_{t+1}^\Lambda$  (noise in the time preference variable) that gets attributed to innovation in stochastic discount factor is the weighted average of  $\lambda_{m,\eta}$  and  $\lambda_{m,\lambda}$  with weights  $\pi_{c,\lambda}$ , and  $\sigma_\Lambda$  respectively. This is because  $\epsilon_{t+1}^\Lambda$  enters both the time preference dynamics as well as the consumption growth process. We can define  $\lambda_{m,x^\Lambda} = [\lambda_{m,\lambda} + \frac{\pi_{c,\lambda}}{\sigma_\Lambda} \lambda_{m,\eta}]$  and rewrite the innovation to stochastic discount factor as

$$m_t - E_t(m_{t+1}) = \lambda_{m,x^\Lambda} \sigma_\Lambda \epsilon_{t+1}^\Lambda + \lambda_{m,\eta} \sigma_{c,t} \eta_{t+1} + \lambda_{m,x} \varphi_e \sigma_{x,t} e_{t+1} + \lambda_{m,\sigma_c} \omega_{c,t+1} + \lambda_{m,\sigma_x} \omega_{x,t+1} \quad (8)$$

$\omega_{c,t+1}$ , and  $\omega_{x,t+1}$  is nothing but the innovation to the stochastic volatility of consumption growth and long run risk process. To be specific,  $\omega_{c,t+1} = E[\sigma_{c,t+1}^2] - \sigma_{c,t+1}^2$ , and  $\omega_{x,t+1} = E[\sigma_{x,t+1}^2] - \sigma_{x,t+1}^2$ . The risk premia can be derived as

$$\begin{aligned}E_t(r_{a,t+1} - r_f) &= 0.5 \text{Var}(r_{a,t+1}) - \text{Cov}[(m_{t+1} - E_t(m_{t+1})), (r_{a,t+1} - E_t(r_{a,t+1}))] + \quad (9) \\ &\quad \Pi[\Omega, R_{a,t+1}] - \Pi[\Omega, R_{f,t+1}]\end{aligned} \quad (10)$$

Substituting the innovation to stochastic discount factor (equation (8)) in risk premia (equation (9)), we get

$$\begin{aligned}E_t(r_{a,t+1} - r_f) + 0.5 \text{Var}(r_{a,t+1}) &= - \left\{ \lambda_{m,x^\Lambda} \text{Cov}(\epsilon_{t+1}^\Lambda, \tilde{r}_{a,t+1}) + \lambda_{m,\eta} \text{Cov}(\sigma_{c,t} \eta_{t+1}, \tilde{r}_{a,t+1}) \right. \\ &\quad + \lambda_{m,x} \text{Cov}(\varphi_e \sigma_{x,t} e_{t+1}, \tilde{r}_{a,t+1}) + \lambda_{m,\sigma_c} \text{Cov}(\omega_{c,t+1}, \tilde{r}_{a,t+1}) \\ &\quad \left. + \lambda_{m,\sigma_x} \text{Cov}(\omega_{x,t+1}, \tilde{r}_{a,t+1}) \right\} + \Pi[\Omega, R_{a,t+1}] - \Pi[\Omega, R_{f,t+1}]\end{aligned} \quad (11)$$

where  $\tilde{r}_{a,t+1} = r_{a,t+1} - E_t(r_{a,t+1})$ . This forms the basic testing equation for cross sectional asset returns. The first stage regression is then given by

$$r_{a,t+1} = \text{cons}_a + \beta_{a,\Lambda} \epsilon_{t+1}^\Lambda + \beta_{a,\eta} \sigma_{c,t} \eta_{t+1} + \beta_{a,x} \varphi_e \sigma_{x,t} e_{t+1} + \beta_{a,\omega_c} \omega_{c,t+1} + \beta_{a,\omega_x} \omega_{x,t+1} \quad (12)$$

where the  $\beta$ 's are the covariance of the corresponding factors with the innovation in asset returns.

### 3 Data Source

The macroeconomic data come from Bureau of Economic Analysis. I take the National Income and Product Accounts (NIPA) tables and construct the growth rate of consumption by taking first difference of log series. The consumption series from NIPA tables are seasonally adjusted so I do not make any further seasonality adjustments. Only the real consumer expenditures on non-durables and real services are considered, consistent with past literature. The data for long run risk measure comes from PROQUEST database. PROQUEST provides access to several newspaper articles from as early as 1925. I access Wall Street Journal as the articles published in The Wall Street Journal (WSJ) are most representative of the US economy. PROQUEST database allows us to enter keywords and obtain all the articles between user-defined data range containing these keywords. The portfolio data comes from Kenneth French's website. I use a total of 45 portfolios to test the model- 25 portfolios sorted by Size and Book-to-market ratio, 10 portfolios sorted by industry, and 10 portfolios sorted by momentum. The standard 25 Size and Book-to-market ratio sorted portfolios lie in low-dimensional excess return space. To improve the statistical power of the Fama-Macbeth regression tests, a wider range of portfolios including industry and momentum sorted portfolios are considered. The real labor income growth and realized industrial production variance is calculated following Liu and Matthies (2016). Real labor income growth is calculated as log differences of real labor income growth from Bureau of Economic Analysis database, and realized industrial production variance is calculated as sum of squared prior 12 months industrial production growth rates from Federal Reserve Economic Data (FRED) database.

## 4 Methodology and Empirical Results

### 4.1 Recovering long run component

The consumption growth process in the model is governed by the below equation

$$g_{t+1} = \mu_c + x_{c,t} + \pi_{c,\lambda} \epsilon_{t+1}^\Lambda + \sigma_{c,t} \eta_{t+1}$$

where  $x_{c,t}$  is the long run risk component that predicts consumption growth. The long run risk component  $x_{c,t}$  is a latent variable that should satisfy two things a) it should reflect the long run news of investor's expectations about the economy b) it should predict the consumption growth, with prediction accuracy increasing over the time horizon. Media news coverage provides a good deal of information about the investor concern regarding the state of economy. Hence we can potentially build an investor concern index based on the news articles, which will proxy for the long run risk component.

Following Liu and Mattheis (2016) [23], I use PROQUEST database that contains historical newspaper article from various sources such as Wall Street Journal, New York Times, Washington

Post, Chicago Tribune etc. I enter the key words "Consumption", and "Gross Domestic Product or GDP" to collect all the daily Wall Street Journal news articles containing these keywords between the years 1925 and 2016. I limit the choice of words to have a higher degree of freedom. Using these keywords gives a total of 61,433 news articles. By glancing through the articles, most of them seem to address the state of economy in the US. For example, one of the articles that shows up in the search is

"U.S. Third Quarter **GDP** Revised Up to 3.5% Gain; But the stronger gains only help bring the year's growth rate back in line. Newly revised data shows consumer spending increased at a 3% annual rate in the summer months, up from an earlier estimate of 2.8% growth, though still a slowdown from the second quarter's robust 4.3% growth rate for household outlays." -Dec 22, 2016."

This article clearly talks about the state of the economy and reflects the investor's concern. However, there are also other articles that contains the keywords but has very little to do with the investor concern. For example, the first few lines of one of the articles that shows up in the search is

"Taste – de gustibus: Election Madness! Cher in Pundit Role, Hollywood in Pain Equally, politicians who run for office are inclined to kiss the celluloid wallet whenever it is proffered – or to snatch at it when it is not. The money that the Clintons have foraged from Hollywood fund-raisers is approaching the **GDP** of Jamaica. -Nov 3, 2000."

This has very little to do with the state of economy. The article shows up because it contains the keyword "GDP" in its text. Inferring stock market movements using such polluted data is a dubious practice, and hence cleaning the data is highly essential. Given that there are a large number of daily news articles, it is time consuming to read each article manually and remove what is not relevant. To circumvent this, I employ several commonly used machine learning models to clean the index through the following steps-

- Take a small random sample of articles from the large database, and classify the articles in the sample as relevant or not relevant manually based on prior knowledge.
- Fit several commonly used machine learning models to this sample
- Pick the best model based on the test error rate
- Use the selected best model to classify the remaining articles in the large database

The first step is to manually classify the articles in the sample as *relevant* or *not relevant*. Thus, the dependent variable is a binary variable with 1 meaning *relevant* and 0 meaning *not relevant*. This

method is called as a "rule-based" approach in machine learning literature where first a manual intervention is employed to classify the data points based on experience or prior knowledge. It is easy to classify articles as *relevant* or *not relevant* in this context by reading the article title and first few lines. The classification will act as dependent variable, while the regressors are the frequency count of words occurring in the article. The next step involves dividing the sample into training observations and test observations. A 70%-30% split is commonly used. The training observations are used to *train* the data, in the sense that several models are fit to this data. The models are built in such a way that it associates frequency of words to the relevance of articles. For example, if a particular word has higher frequency in the articles that are flagged as 1, then this word is given a higher score such that the probability that an article containing this word is *relevant* is large. On the other hand, if a particular word repeatedly occurs in the articles that are flagged as 0, then this word is given a lower score such that probability of an article containing this particular word being *relevant* is small. Once the models are fit in this fashion, we take the fitted models and apply it to the test data and measure the performance of each of these models. Although the test data also contains manual classification of flags (0 and 1), we pretend that this assignment does not exist and we ask the fitted model to classify it for us based on the prediction model that has been built. The classification that the model gives us is called as predicted classification. We then compare it to the true classification, where true classification is the manual classification done in the first step. The total number of misclassifications is called as the *test error rate*. The model that gives the lowest *test error rate* is chosen to be the best model for the sample at hand. It is important to have a large sample for the training data otherwise the test data may have several words with no score attached to it simply because these words did not occur in the training data. This procedure is used in Lee and Lee (2008) [21] to predict the 8-K content from corporate filings.

I take a random sample of 2,500 articles from 61,433 articles and read the title and abstract for each of them. If they are relevant to the state of economy, I flag it as 1, and 0 otherwise. I clean the text in the articles by removing the punctuation marks and numbers, removing the commoner morphological and inflexional endings using Porter Stemming Algorithm, and converting all the text to lower case. I then create a frequency matrix containing the words as columns, and the articles as rows. The matrix size is 2,500 X 10,905. The cells take a value  $n \in \mathbb{N}^0$  where  $n$  is the number of times the particular word in column occurs in the corresponding article in the row. The figure (7) in Appendix A shows the words that occur at least a 150 times in the articles considered in the sample. Words like "market", "year", "price", etc., as expected occur quite frequently. There are many words that occur in very few articles making the frequency matrix very sparse. In fact, only 0.33% cells (90989/27171511) of the frequency matrix take non zero values. Hence, I remove the sparse words by dropping from the frequency matrix the words which have at least 99.5% percentage of empty elements. The resulting matrix has size 2,500 X 1,344. A significant reduction in the number of words from 10,905 to 1,344 has occurred. I take 70% of the sample

(1,667 articles) as *training observations* and the remaining 30% (833 articles) as *test observations*. I adopt commonly used machine learning tools such as Logistic regression, Ridge regression, LASSO regression, Classification tree, Support Vector Machines, Bagging, and Random Forests. I find that Random forest provides the lowest test error rate of 19%<sup>10</sup>. I classify the remaining 58,933 (61,433 -2,500) articles as *relevant* or *not relevant* using the random forest algorithm. In this way, we would have all the articles in the database classified. Finally, I remove the articles that have flag 0 from the database since they are deemed *not relevant*. Among the *relevant* articles, I count the number of articles in each month to create the monthly news index. I detrend the news index and normalize it to have a standard deviation of 1. The figure 1 shows the final news index measure. The shaded regions indicate the NBER recession dates.

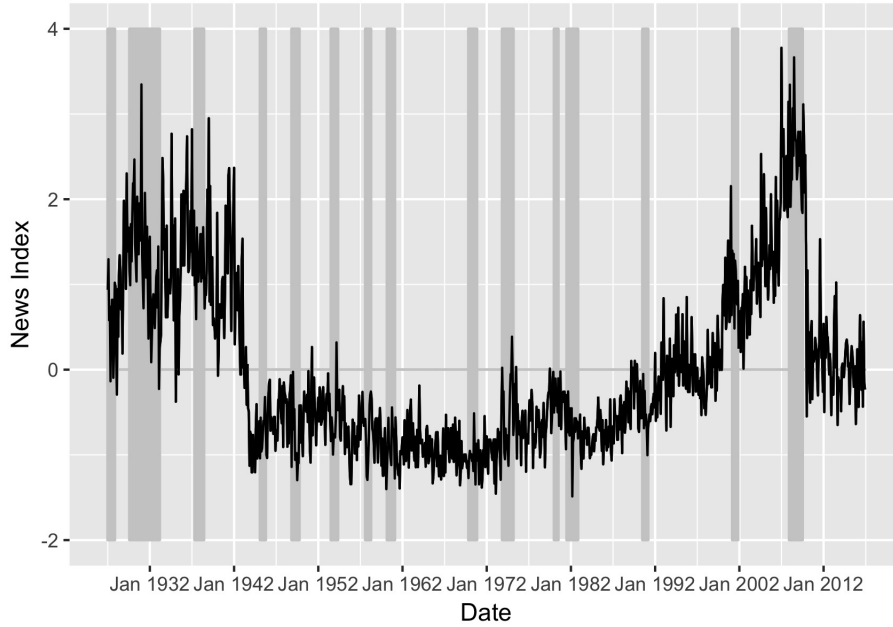


Figure 1: News Index

The index takes a high value in times of economic crisis with periods in the range 1930-1940 and 2008-2012 having peak investor concern. This is because media tends to pay more attention to the economy in times of distress.

#### 4.1.1 Predictability

The predictability of consumption growth is extremely important in the Macro-Finance literature. Bansal-Yaron long run risk model starts from the assumption that there *is* a predictable component in consumption growth. Other models such as Inter-temporal capital asset pricing model (ICAPM- Merton, 1973)[25], and Habit (Campbell and Cochrane 1999 [8]) use a random walk consumption growth as postulated by Hall (1988) [16]. The rare disaster risk model by Rietz

<sup>10</sup>The summary of performance results of various classification algorithm is provided in table (3) in Appendix A

(1988) [31] and Barro (2006) [5] states that while there may be a predictable component in the consumption process, it will be hard to estimate it empirically especially in finite sample. Bansal et al (2014)[4] show that Industrial Production Variance, and Labor Income Growth can predict the short term consumption growth, while Lettau and Ludvigson (2001) [19] demonstrate that consumption wealth ratio (CAY) can predict the stock returns in future. Gourio (2012) [15] uses a time varying jump probability in the consumption growth to estimate the predictability. Several approaches have been used in the predictability literature in the past with several variables shown as having predictability power. The econometrics literature on predictability show that the nature of predictability regression induces a bias in the coefficients since these regressions involve several overlapping samples. Newey and West (1987) [27] provide an error correction method to tackle this issue. Hence, all return predictability regressions should incorporate the Newey West adjusted standard errors.

The news index should predict the future consumption growth over the longer term since it acts as a proxy for the long run risk. The consumption growth predictability equation is given by

$$\sum_{i=1}^K g_{t+i} = \beta_0 + \beta_1 \text{NewsIndex}_t + \epsilon_{t+K}, \quad K = 1, 2, \dots, 100$$

where  $K$  indicates the months and  $g_{t+i}$  represents the consumption growth between periods  $t+i$  and  $t+i-1$ . Since the news measure is a long run risk measure, the predictability is expected to increase over the time horizon. The results for the predictability regression confirm this hypothesis. The figure (8) in Appendix A plots the adjusted R-squared for 100 separate regressions for 100 rolling time horizons (months). For example, the first regression is between  $g_{t+1}$  and  $\text{NewsIndex}_t$ , second regression is between  $(g_{t+1} + g_{t+2})$  and  $\text{Index}_t$  and so on. The adjusted R square increases over the time horizon taking a maximum value of 57% at 86th month. This means that the predictability of consumption growth is strongest at 86 months, roughly around 7 years. The predictability power declines after this period. The slope coefficient is higher for longer time periods than the shorter time periods and is statistically significant at 1% level even after correcting standard errors using Newey West error correction method. As Liu and Matthies (2016)[23] point out, the real labor income growth, price dividend ratio, and realized industrial production variance can affect the consumption growth process. Hence, I also perform multivariate regressions with these three variables as control variables. The results from multivariate regression (see table (5) in Appendix A) indicate that the coefficients of long run risk measure remain close to the coefficients from simple regression, and is significant at 1% level. Therefore, the news index is robust to inclusion of other variables when it comes to explaining the long run aggregate consumption growth. In other words, the additional variables do not add much predictive power on top of the news index.

## 4.2 Recovering other state variables

The time preference shock is latent too and needs to be recovered. Recall the consumption growth process,

$$g_{t+1} = \mu_c + x_{c,t} + \pi_{c,\lambda} \epsilon_{t+1}^\Lambda + \sigma_{c,t} \eta_{t+1}$$

A scaled up innovation to the time preference shock process ( $\sigma_\Lambda \epsilon_{t+1}^\Lambda$ ) enters the consumption growth process. Although we don't directly observe the time preference risk (innovation to time preference shock), we can see that it is linear in risk free rate and price dividend ratio <sup>11</sup>. Define a new variable  $\tilde{g}_{t+1} = g_{t+1} - x_{c,t}$ . The time preference risk can then be recovered from regression of  $\tilde{g}_{t+1}$  on price dividend ratio and risk free rate. Specifically,  $\tilde{g}_{t+1} = \mathbf{b}'\mathbf{Y}_t + \sigma_t \eta_{t+1}$  where  $\mathbf{Y}_t = [1 \quad z_{m,t} \quad r_{f,t}]$ . The time preference risk is then given by  $\mathbf{b}'\mathbf{Y}_t$ . However,  $\tilde{g}_{t+1}$  is highly non-stationary which implies OLS estimation would result in a spurious regression. This calls for the need for a more efficient way of extracting the innovation to time preference variable. I make use of the Hodrick Prescott Band pass filter which delineates the trend component from the cyclical component of  $\tilde{g}_{t+1}$ . Hodrick-Prescott band pass filter is documented to have a superior performance among the other commonly used band pass filters such as Christiano-Fitzgerland, and Baxter-King filter (Nilsson and Gyomai, 2011) [28]. Since the time preference variable relates to the long run expectations of investors, a smoothed representation which is more sensitive to the long run changes rather than short term fluctuations is relevant. The Hodrick Prescott filter optimizes

$$\min_{\tau} \left( \sum_{t=1}^T (y_t - \tau_t)^2 + \lambda \sum_{t=2}^{T-1} [(\tau_{t+1} - \tau_t) - (\tau_t - \tau_{t-1})]^2 \right)$$

where  $y_t$  is the time series data comprising of the trend  $\tau$  and cyclical component  $c$ . The first term  $(y_t - \tau_t)$  refers to the sum of squares of cyclical component, and the second term  $\lambda \sum_{t=2}^{T-1} [(\tau_{t+1} - \tau_t) - (\tau_t - \tau_{t-1})]^2$  refers to the sum of squared of second differenced trend component multiplied by a penalty factor  $\lambda$ . HP filter is applicable only when the data is second difference stationary. Since  $\tilde{g}_{t+1}$  (consumption growth minus long run component) is second difference stationary, HP filter can be applied. I choose  $\lambda = 129600$  for monthly data as suggested by Ravn and Uhlig (2002) [30]. Note that a higher value of  $\lambda$  increases the penalty factor and hence variation in growth rate of trend component is heavily penalized. The time preference risk is displayed in figure (2). As expected, the risk is the highest around the period 2007-2009 during the times of financial crisis. The investors like to smooth their time preference shock over time. Hence, a higher innovation to the time preference shock is disliked by the investors. This innovation to the time preference shock is what I call as time preference risk. In the perspective of risk premia estimation equations, what matters is the time preference risk and not the time preference shock per se because what enters the consumption growth process is the innovation to the time preference shock. While the trend component from applying the filter corresponds to the time preference shock, the residual component corresponds to the noise in consumption process. This consumption

---

<sup>11</sup>This is because price dividend ratio follows the process  $z_{m,t+1} = A_{0,m} + A_{1,m}x_{c,t+1} + A_{2,m}x_{t+1}^\Lambda + A_3\sigma_{c,t}^2 + A_4\sigma_{x,t}^2$



noise is nothing but the short run risk. I plot the innovation to the time preference shock (time preference risk) in figure (2). Figure (3) provides the consumption growth, and innovation to the consumption growth process or in other words short run risk. I also superimpose the consumption growth against the recovered persistent components in figure (9) in Appendix A. We can see that the state variables track the consumption growth process closely.

The next step is to extract the long run risk, which is nothing but the noise in the dynamic equation of long run component. Recall that the long run persistent component follows

$$x_{c,t+1} = \rho x_t + \varphi_e \sigma_t e_{t+1}$$

We could fit an AR(1) process and then recover the noise from the AR(1) process. An alternative would be to use a band pass filter since they are known to be efficient in extracting noise from the signal in time series data. In particular, the band pass filters provide a smoothed estimate of the state variables. The innovation in the news measure extracted using Hodrick-Prescott filter is shown in figure (2). The innovation is higher during the periods of economic distress which is intuitive because shocks to the long run risk component occur when economic prospects are bleak.

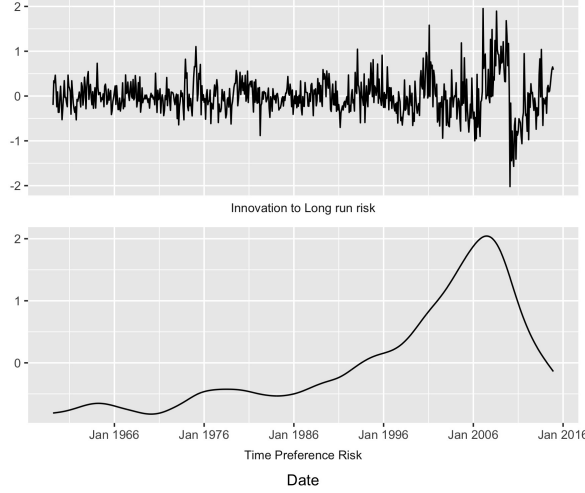


Figure 2: State variables

Lastly, the innovation to volatility is recovered by fitting an ARIMA(0,1,2) process to the extracted volatility variable. Several configuration of ARIMA models are fit to the volatility state and the best model is selected based on AIC and BIC. ARIMA(0,1,2) implies that the volatility follows an exponentially weighted moving average process. It turns out that the time varying component in consumption growth volatility ( $\sigma_{c,t}$ ) and the time varying component in long run volatility ( $\sigma_{x,t}$ ) has a correlation of around 97%. In the original Bansal-Yaron (2004) model these two components are the same. That is, the same time varying component enters the stochastic volatility of consumption growth as well as long run component. Hence, to retain parsimony in the model, I assume  $\sigma_{c,t} = \sigma_{x,t}$ , which is not unreasonable given the high correlation post estimation.

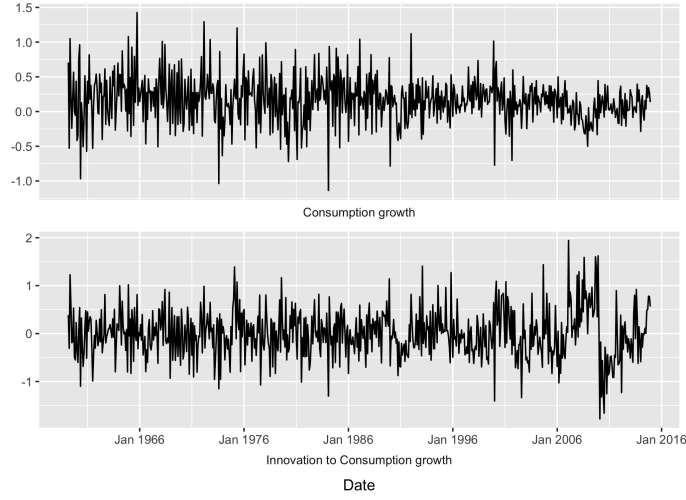


Figure 3: Consumption process

The implication of this is that the number of state variables reduces to three ( $\mathbf{S}_t = [x_{c,t} \ x_t^\Lambda \ \sigma_{c,t}^2]$ ) from four ( $\mathbf{S}_t = [x_{c,t} \ x_t^\Lambda \ \sigma_{x,t}^2 \ \sigma_{c,t}^2]$ ). Also, the volatility risk term  $\sigma_{x,t}^2$  disappears from the asset price testing equations.

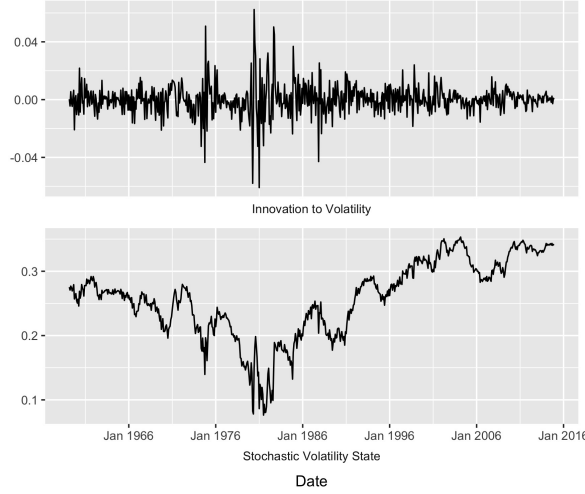


Figure 4: Volatility State

### 4.3 Measurement Error

Before moving on to the estimation of risk premia, I study the magnitude of the bias term that appears in the Euler equation. If the bias is non-zero, then the estimation of risk premia should be such that the bias is taken into consideration. The bias term is given by the expression  $\Pi \left[ \Omega, R_{i,t+1} \right]$ , which is a function of utility parameters (risk aversion  $\gamma$ , inter-temporal elasticity of substitution  $\psi$ ), discount rate ( $\delta$ ), time preference shock  $\frac{\lambda_{t+1}}{\lambda_t}$ , the aggregate wealth portfolio return ( $R_{a,t+1}$ ), the asset return ( $R_{i,t+1}$ ), and the variance of measurement error ( $\sigma_\epsilon^2$ ). To provide an estimate for

the bias, I take range of plausible values for the utility parameters such that the range includes the estimated values from past literature. For example, the estimated values for risk aversion, inter-temporal elasticity of substitution and discount factor are 8.6, 1.93, and 0.999 respectively in Schorfheide et al (2016) [33]. Hence, I take a range of values for these parameters such that the interval includes these estimates. I assume two different measurement error variance ( $\sigma_\epsilon$ )- 0.005%, and 0.05%. These are fairly low numbers which will give us a conservative estimate of the bias. I use the term "measurement error" here to refer to the measurement error in Euler equation induced by the measurement error in consumption data. I simulate the time preference shock ( $\frac{\lambda_{t+1}}{\lambda_t}$ ) using the estimates from Schorfheide et al (2016) [33]. The value weighted return of all CRSP (The Centre for Research in Security prices) firms incorporated in the US and listed on the NYSE (New York Stock Exchange), AMEX (American Stock Exchange), or NASDAQ (National Association of Securities Dealers Automated Quotations) is used as proxy for return on aggregate wealth portfolio. The bias is computed for return on market portfolio ( $R_{i,t+1}$  in the bias term becomes  $R_{m,t+1}$ ).

A summary of the measurement error is provided in Table (1). Even with a measurement error variance of 0.005% which is considerably small (Panel A), the measurement error in Euler equation is non-zero ranging from 2.2% to 8.1%. The error increases along with risk aversion and decreases along with IES. For a measurement error variance of 0.05%, which is still small, the bias shoots up ranging from 22% to 75%. This demonstrates the importance of measurement error in consumption data since the bias in the Euler equation seems to be large even for moderate values of measurement error variance. The Euler equation is the same as the pricing equation and hence this bias can be seen as pricing errors. It is hard to rely on a pricing model with such high pricing errors. Evidently, the problem of measurement error cannot be ignored. Figures (10) and (11) in Appendix A provide a visualisation of the measurement error for various values of utility parameters and error variance.

#### 4.4 Risk premia estimation

I use the model from Giglio and Xiu (2017) [14] since a three pass regression procedure would achieve robust risk premia estimation in the presence of measurement error or unobserved heterogeneity. Consider the vector of observable macro factors

$$g_t = \left[ \text{time preference risk, long run risk, short run risk, volatility risk} \right]$$

Due to the measurement error, the second stage Fama MacBeth regression contains omitted factors. Consider the pricing model

$$r_t = \iota_n \gamma_0 + \alpha + \beta \gamma + \beta \nu_t + u_t \quad f_t = \mu + \nu_t \quad E(\nu_t) = E(u_t) = 0 \quad Cov(u_t, \nu_t) = 0 \quad (13)$$

where  $\nu_t$  is a  $p \times 1$  vector of innovations of  $f_t$ ,  $u_t$  is a  $n \times 1$  vector of pricing errors,  $\beta$  is an  $n \times p$  factor loading matrix, and  $\gamma_0$  and  $\gamma$  are zero-beta rate and  $p \times 1$  risk premia vector. There are  $p$  factors in the model among which  $g$  factors are observed. The observable factor returns are given

Risk Aversion						Risk Aversion					
Panel A						Panel B					
	7.6	8	8.5	9			7.6	8	8.5	9	
IES	1.81	3.6%	4.2%	5.1%	6.1%	IES	1.81	36%	42%	51%	61%
	1.85	3.2%	3.7%	4.5%	5.4%		1.85	32%	37%	45%	54%
	1.9	2.8%	3.3%	3.9%	4.7%		1.9	28%	33%	39%	47%
	1.95	2.5%	2.9%	3.4%	4.1%		1.95	25%	29%	34%	41%
	2	2.2%	2.6%	3.1%	3.6%		2	22%	26%	31%	36%
Risk Aversion						Risk Aversion					
Panel C						Panel D					
	7.6	8	8.5	9			7.6	8	8.5	9	
IES	1.81	3.8%	4.4%	5.4%	6.6%	IES	1.81	38%	44%	54%	66%
	1.85	3.4%	4.0%	4.8%	5.8%		1.85	34%	40%	48%	58%
	1.9	3.0%	3.5%	4.2%	5.0%		1.9	30%	35%	42%	50%
	1.95	2.6%	3.1%	3.7%	4.4%		1.95	26%	31%	37%	44%
	2	2.2%	2.6%	3.1%	3.6%		2	23%	27%	32%	39%

Table 1: Measurement Error Summary. Panel A: Measurement error variance 0.005%, discount factor 0.999. Panel B: Measurement error variance 0.05%, discount factor 0.999. Panel C: Measurement error variance 0.005%, discount factor 0.995. Panel D: Measurement error variance 0.05%, discount factor 0.995.

by

$$g_t = \xi + \eta\nu_t + z_t \quad E(z_t) = 0, \quad Cov(z_t, \nu_t) = 0$$

where  $n \times p$  matrix  $\eta$  is the loading of the factor  $g$  on  $\nu$ , and  $z_t$  is the measurement error. Note that once we remove measurement error, the observable factor  $g_t$  is a linear transformation of fundamental factors  $\eta\nu_t$ . Thus, the true asset pricing model comprises  $g_t$  which corresponds to the first  $d$  factors, and remaining  $p-d$  factors correspond to linear combination of unobservable factors. Our goal is to obtain the robust risk premia estimate of each of factor in the vector  $g_t$ . This risk premia is given by  $\eta\gamma$  and as long as the space spanned by the rotated factors is same as that of true factors, this premia is invariant to whatever factors that we choose in model (13). This leads to the three pass estimation procedure.

### Three pass estimator

- Step 1 (PCA). Extract the principal components of returns by conducting principal component of covariance matrix of returns. The covariance matrix is scaled by  $T^{-1}n^{-1}$  since it is shown to accelerate the matrix decomposition algorithm when  $T > n$ . The estimator for factor scores and factor loadings are then given by

$$\hat{\beta} = T^{1/2}(\varepsilon_1|\varepsilon_2|.....|\varepsilon_p)$$

where  $(\varepsilon_1|\varepsilon_2|.....|\varepsilon_{\hat{p}})$  is the eigenvectors corresponding to largest  $\hat{p}$  eigenvalues of covariance

matrix of returns. The factor scores are given by

$$\hat{V} = T^{-1} \bar{R} \hat{V}^T$$

where  $\bar{R}$  is the demeaned time series of asset returns.

- Step 2 (CSR): Obtain the risk premia of estimated factor scores by running a cross sectional regression of average asset returns onto estimated factor loadings  $\hat{\beta}$ . The risk premia of the estimated factor returns is given by

$$\tilde{\Gamma} := (\tilde{\gamma}_0, \tilde{\gamma}^T)^T = \left( (\iota_n : \hat{\beta})^T (\iota_n : \beta) \right)^{-1} (\iota_n : \hat{\beta})^T \bar{r}$$

- Step 3 (TSR): Obtain the exposure of each observed factor to estimated factor scores. This is obtained by running a time series regression of observed factors against the estimated factor scores.

$$\hat{\eta} = \bar{G} \hat{V}^T (\hat{V} \hat{V}^T)^{-1}, \quad \text{and} \quad \hat{G} = \eta \hat{V}$$

The risk premia is given by

$$\hat{\Gamma} := \begin{pmatrix} \hat{\gamma}_0 \\ \hat{\gamma} \end{pmatrix} := \begin{pmatrix} 1 & 0 \\ 0 & \hat{\eta} \end{pmatrix} = \begin{pmatrix} \tilde{\gamma}_0 \\ \hat{\eta} \tilde{\gamma} \end{pmatrix}$$

The optimal number of eigenvectors  $\hat{p}$  in the first step can be chosen from the magnitude of eigenvalues. Figure (5) plots the first 10 eigenvectors of the covariance matrix of returns. In large panels of asset returns, the first eigenvalue typically tends to be much larger than the others. Hence, I plot the eigenvalues excluding the first one on the right panel of Figure 8. After 4 and 7 factors, there is a noticeable decrease in the eigenvalues. The three pass approach is consistent as long as  $\hat{p}$  is at least as large as the true panel dimension. Hence, I choose  $\hat{p} = 7$ .

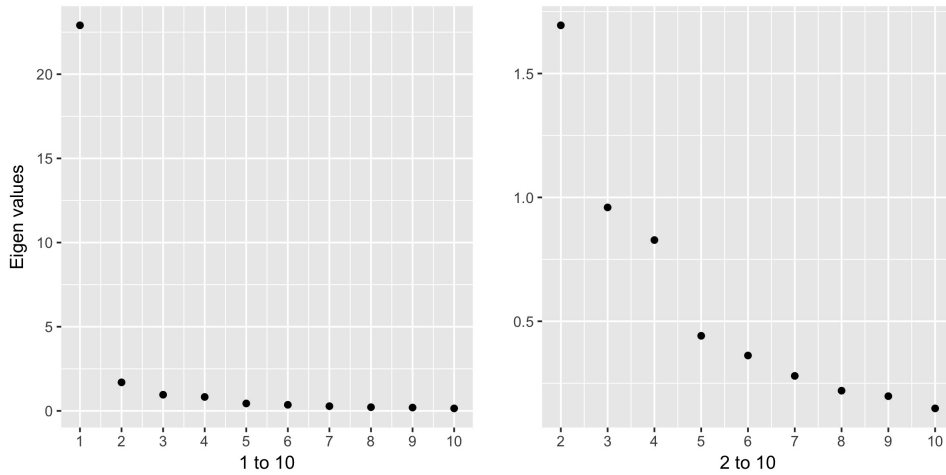


Figure 5: Scree plot

### Three pass estimation results

- I extract the first two principal components from the covariance matrix of returns scaled by  $T^{-1}n^{-1}$  and plot it along with the eigenvectors in Figure (6) . It can be seen that the first principal component explains about 87% variation in the asset returns and the second principal component explains about 6% variation. The latent factors that drive the asset returns seem to be strongly correlated to almost all of the stocks with a possible exception of utility stocks.

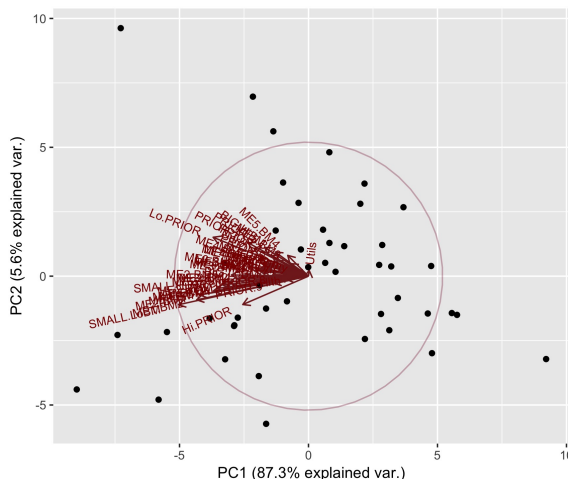


Figure 6: Principal Component Analysis

- The cross sectional regression of asset returns on the estimated factor scores provides an adjusted R-squared of 75%. Such high R-squared is not surprising since the factor scores are obtained from the statistical decomposition of the asset returns in the first place. The regression result is given in Table (6) in Appendix A.
- The time series regression of the individual factors on the estimated factor scores provide a very low fit (less than 5%). This is in line with Giglio and Xue (2017) [14] who find that the macroeconomic variables have a weak relation to the latent factors recovered from the principal component analysis. This essentially implies that the macro factors are not related to the underlying statistical latent factor that is recovered using PCA. Combining step 2 and steps 3, the risk premia estimates are obtained and summarized in table (2). Comparison to risk premia estimates from regular two pass Fama MacBeth procedure is also given in table (2).

The close to zero risk premia estimates and very low R-squared in the third stage provide a strong evidence *against* the long run risk and time preference risk model. This is in contrast to strong evidence in favour of the long run risk and time preference risk model using the usual Fama MacBeth two pass estimation. This difference can be attributed to the problem of omitted factors.

Table 2: Risk premia estimates

Risk Premia estimates	3 pass method	Fama MacBeth
Preference risk	0.0052	0.561109**
Long run risk	0.0029	0.285747**
Short run risk	-0.0009	0.273237*
Volatility risk	0.00037	0.007335***

\*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.10$

which stems from the measurement error in consumption data in this case.

Cochrane (2016) [9] points out that long run risk models have a deep connection to the early resolution of uncertainty. Recursive utility requires that consumers prefer to resolve the uncertainty earlier. What long run risk model states is that the investors like to know about the uncertainty in future even though they cannot provide a response to it in the present. It is hard to believe that people would like to receive news about the future and when they do they will not act upon it now. Cochrane (2016) [9] makes an interesting analogy to the Huntington’s disease for which there is no cure. A test for this genetic disease reveals whether people have the disease or not but unfortunately a cure doesn’t exist. As results from Oster et al (2013) [29] indicate, people with family history of Huntington’s disease rarely get the test. This is a news which they can acquire and possibly even provide a response to it in terms of altering the financing choices and personal lifestyle choices. Yet, very few people take this test which questions the very hypothesis of early resolution of uncertainty.

## 5 Conclusion

I have used news articles from The Wall Street Journal to create the long run risk measure employing machine learning techniques. This news index predicts long run aggregate consumption growth 7 years down the line with an adjusted R-squared of 57% and is robust to inclusion of other commonly used variables in predictability literature. Measurement error in consumption data is modelled using an additive error model and an adjusted Euler equation is derived to take into account the measurement error. The bias induced from this measurement error in the pricing equation is non-zero, even if we assume a very low measurement error variance. The modified risk-premia estimation equations indicate that the bias enters the second stage regression equation as an additional unobservable component. The results from the three pass estimation technique proposed by Giglio and Xiu (2017) [14] indicate that the robust risk premia estimates of the long run risk factor and time preference factor are close to zero, and these factors are spurious. This is in contrast to the result from regular two-pass estimation technique which indicate that these factors have non-zero risk premia and explain the cross section of asset returns. The model therefore does

not gain support when using *robust* estimation technique.

This paper contributes to the recent literature of long run risk models by using text based information from the media to infer long run investor expectations. Media content is usually noisy but given the increasing success of machine learning tools in prediction problems, one can proxy the latent components such as unobserved risk measures with a good approximation. This opens up an avenue to research in estimating unobservables in economic models using textual information. The second contribution comes from theoretically deriving an estimatable bias in adjusted Euler equation of the model. Lastly, while the prior literature to test long run risk models have used standard Fama MacBeth two pass approach or GMM, this paper is the first to use a three pass estimation technique to obtain robust risk premia estimates.

Few extensions could be made to this paper. The news index is created using a "rule-based" approach where manual intervention is needed to classify the sample articles as *relevant* or *not relevant*. This could be replaced by a purely statistical approach where the model can be built in such a way that the frequency of words predict whether the news article is related to the consumption growth or not. Also, the news index takes peak values in great depression and global financial crisis but not during periods of war or political shocks. Muir (2016)[[26](#)] points out that during periods of war, the fluctuation in consumption is high but movement in asset prices are low. On the other hand, during periods of financial crisis, the fluctuation in consumption is low but movement in asset prices are high. While the news index captures the sentiment during financial crisis, one could build an index that also captures sentiment during periods of war and political shocks in order to better explain stock price movements.



## References

- [1] Albuquerque, R. A., M. Eichenbaum, V. X. Luo, and S. T. Rebelo (2015): "Valuation risk and asset pricing", *Journal of Finance*, Forthcoming.
- [2] Bansal, R., D. Kiku, and A. Yaron (2012): "An empirical evaluation of the long-run risks model for asset prices", *Critical Finance Review*, 1, 183-221.
- [3] Bansal, R., and A. Yaron (2004): "Risks for the long run: A potential resolution of asset pricing puzzles", *Journal of Finance*, 59.4, 1481-1509.
- [4] Bansal, R., D. Kiku, I. Shaliastovich, and A. Yaron (2014): "Volatility, the Macroeconomy, and Asset Prices", *Journal of Finance*, 69, 2471-2511
- [5] Barro, R. J. (2006): "Rare disasters and asset markets in the twentieth century", *Quarterly Journal of Economics*, 121, 823-866.
- [6] Brav, Alon, George M Constantinides, and Christopher C Geczy (2002): " Asset pricing with heterogeneous consumers and limited participation: Empirical evidence", Technical report, National bureau of economic research.
- [7] Campbell, J. Y., and R. J. Shiller (1988): "The dividend-price ratio and expectations of future dividends and discount factors", *Review of Financial Studies*, 1.3, 195-228.
- [8] Campbell, J.Y., and Cochrane, J. (1999): "By force of habit: A consumption based explanation of aggregate stock market behaviour", *Journal of Political Economy*, Volume 107, Issue 2 (April 1999), 205-251
- [9] Cochrane, John (2016): "Macro-Finance", NBER Working paper.
- [10] Da, Zhi, and Hayong Yun (2010): "Electricity consumption and asset prices", Available at SSRN 1608382 .
- [11] De, K (2015), "Consumption Based Asset Pricing Adjusting for Measurement Error in Consumption", Available at SSRN.
- [12] Epstein, Larry G, and Stanley E Zin (1989): "Substitution, risk aversion, and the temporal behavior of consumption and asset returns: A theoretical framework", *Econometrica: Journal of the Econometric Society* 937-969.
- [13] Fama, E. F., and J. D. MacBeth (1973): "Risk, return, and equilibrium: Empirical tests", *Journal of Political Economy*, 81, 607-636.
- [14] Giglio, S, and Xiu, D (2017): "Inference on Risk Premia in the Presence of Omitted Factors", Working Paper, Chicago Booth.
- [15] Gourio, F. (2012): "Disaster Risk and Business Cycles", *American Economic Review*, 102(6), 2734-66.

- [16] Hall, R. (1988), "Intertemporal Substitution in Consumption", *Journal of Political Economy*, April 1988, 96(2), pp. 339-357.
- [17] Hansen, L. P., and K. J. Singleton (1982): "Generalized instrumental variables estimation of nonlinear rational expectations models", *Econometrica*, 50, 1269-1286.
- [18] Hodrick, Robert J. and Edward C. Prescott. 1997. "Postwar U.S. Business Cycles: An Empirical Investigation", *Journal of Money, Credit and Banking*, 29(1): 1-16
- [19] Lettau, M., and S. Ludvigson (2001a): "Resurrecting the (C) CAPM: A cross-sectional test when risk premia are time-varying", *Journal of Political Economy*, 109.6, 1238-1287.
- [20] Lewellen, J., S. Nagel, and J. Shanken (2010): "A skeptical appraisal of asset pricing tests", *Journal of Financial Economics*, 96.2, 175-194.
- [21] Lee, J.M., and Lee, H. (2008), "Predicting Corporate 8-K Content Using Machine Learning Techniques", Graduate School of Business Stanford University.
- [22] Li, Feng (2010), "The Information Content of Forward-Looking Statements in Corporate Filings ? A Naive Bayesian Machine Learning Approach", *Journal of Accounting Research*, Vol. 48 No. 5.
- [23] Liu, L., and Matthies, B., (2016): "Long Run Risk: Is It There?", NBER Working paper.
- [24] Mehra, R. (2003): "The equity premium: Why is it a puzzle?", NBER Working paper.
- [25] Merton, Robert C., (1973): "An intertemporal capital asset pricing model", *Econometrica* 41, 867-887
- [26] Muir, T. (2016): "Financial Crises and Risk Premia", Available at SSRN. (Forthcoming in *Quarterly Journal of Economics*).
- [27] Newey, W. K., and K. D. West (1987): "A simple, positive semi-definite, heteroskedasticity and autocorrelation and consistent covariance matrix", *Econometrica*, 55.3, 703-708.
- [28] Nilsson, R. and Gyomai, G., (2011): "Cycle Extraction: A Comparison of the Phase-Average Trend Method, the Hodrick-Prescott and Christiano-Fitzgerald Filters," *OECD Statistics Working Papers* 2011/4, OECD Publishing.
- [29] Oster, E., Shoulson, I., and Dorsey, E.R., (2013): "Optimal Expectations and Limited Medical Testing: Evidence from Huntington Disease?", *American Economic Review* 103:804-813
- [30] Ravn, O.M., and Uhlig, H. (2002): "On adjusting the Hodrick-Prescott filter for the frequency of observations", *The Review of Economics and Statistics*, 84 (2) pp. 371-376.
- [31] Rietz, T., (1988): "The Equity Risk Premium: A Solution", *Journal of Monetary Economics*, 22, 117-131.

- [32] Savov, A. (2011): "Asset pricing with garbage", *Journal of Finance*, 66, 177-201.
- [33] Schorfheide, F., D. Song, and A. Yaron (2016): "Identifying long-run risks: A bayesian mixed-frequency approach", Working Paper.
- [34] Vissing-Jorgensen, Annette (1998): "Limited stock market participation and the equity premium puzzle".

## A Appendix A

### A.1 Model

A representative agent model with endowment economy and preferences characterized by Epstein and Zin (1989) is considered. The agent, holding recursive preferences, maximises the lifetime utility by solving the following Hamilton Jacobi Bellman equation

$$U_t = \max_{C_t} \left\{ \lambda_t C_t^{1-\frac{1}{\psi}} + \delta ([U_{t+1}^*])^{1-\frac{1}{\psi}} \right\}^{\frac{1}{1-\frac{1}{\psi}}} \quad (14)$$

subject to the budget constraint,

$$W_{t+1} = (W_t - C_t)R_{c,t+1} \quad (15)$$

$C_t$  is the consumption stream,  $\psi$  is the IES,  $\theta = \frac{1-\gamma}{1-\frac{1}{\psi}}$ , and  $\delta$  refers to the agent's time preference.  $R_{c,t+1}$  is the return on consumption claim,  $U_{t+1}^*$  is the certainty equivalent value of lifetime utility from  $t+1$  such that

$$(U_{t+1}^*)^{1-\gamma} = E_t(U(W_{t+1})^{1-\gamma})$$

The risk aversion parameter is given by  $\gamma$ . When  $\gamma = 1/\psi$ , the recursive preferences reduces to CRRA preferences. The demand shocks is captured by the changes in time discount rate of investors. The variable  $\Lambda_{t+1} = \lambda_{t+1}/\lambda_t$  captures how the investors trade off current utility versus future period utility. Let  $x_{\Lambda,t}$  denote the time preference shock, which follows the law of motion

$$x_{\Lambda,t+1} = \rho_{\Lambda} x_{\Lambda,t} + \sigma_{\Lambda} \epsilon_{t+1}^{\Lambda}$$

where  $\epsilon_{t+1}^{\Lambda} \sim N(0,1)$ . The law of motion for consumption process is a modified version of Bansal-Yaron type since the innovation to the time preference shock enters the consumption growth process. The state process is given by

$$\begin{aligned} g_{t+1} &= \mu_c + x_{c,t} + \pi_{c,\lambda} \epsilon_{t+1}^{\Lambda} + \sigma_{c,t} \eta_{t+1} \\ x_{c,t+1} &= \rho x_t + \varphi_e \sigma_{x,t} e_{t+1} \\ \ln \frac{\lambda_{t+1}}{\lambda_t} &= \rho_{\Lambda} \ln \frac{\lambda_t}{\lambda_{t-1}} + \sigma_{\Lambda} \epsilon_{t+1}^{\Lambda} \\ \sigma_{k,t+1}^2 &= \nu + \sigma_{k,t}^2 + \phi_1 \zeta_t - \phi_2 \zeta_{t-1} - \phi_3 \zeta_{t-2}, \quad k \in \{c, x\} \\ \eta_t, e_t, \epsilon_t, \zeta_t &\sim \text{i.i.d standard-normal random variable} \end{aligned}$$

Using the log-linearization of the return formula as in Campbell and Shiller (1988), we have the return of an asset in the form

$$r_{a,t+1} = \kappa_0 + \kappa_1 z_{t+1} - z_t + g_{t+1}$$

where  $z_t$  is defined as the log price consumption ratio. A conjecture for  $z_t$  is

$$z_t = A_0 + A_1 x_{c,t+1} + A_2 x_{t+1}^{\Lambda} + A_2 \sigma_{c,t+1}^2 + A_4 \sigma_{x,t+1}^2$$

Similarly, the return on market portfolio and price dividend ratio is given by

$$r_{m,t+1} = \kappa_{0,m} + \kappa_{1,m} z_{m,t+1} - z_{m,t} + g_{d,t+1}$$

where  $g_{d,t+1}$  is the dividend yield growth. The price dividend ratio for market takes the affine form

$$z_{m,t} = A_{m,0} + A_{m,1} x_{c,t+1} + A_{m,2} x_{t+1}^\Lambda + A_{m,3} \sigma_{c,t+1}^2 + A_{m,4} \sigma_{x,t+1}^2$$

## A.2 Solving for stochastic discount factor

The first order condition from the Bellman equation is given by

$$FOC \quad C_t | \quad \lambda_t C_t^{-\frac{1}{\psi}} = \delta (U_{t+1}^*)^{-\frac{1}{\psi}} \left[ E_t \left( U(W_{t+1})^{1-\gamma} \right) \right]^{\frac{1}{1-\gamma}-1} E_t \left( U(W_{t+1})^{-\gamma} U'(W_{t+1}) R_{c,t+1} \right) \quad (16)$$

Envelope condition is given by

$$U'(W_t) = U(W_t)^{-\frac{1}{\psi}} \delta (U_{t+1}^*)^{-\frac{1}{\psi}} \left[ E_t \left( U(W_{t+1})^{1-\gamma} \right) \right]^{\frac{1}{1-\gamma}-1} E_t \left( U(W_{t+1})^{-\gamma} U'(W_{t+1}) R_{c,t+1} \right) \quad (17)$$

Dividing equations (16) and (17), we get

$$\frac{\lambda_t C_t^{-\frac{1}{\psi}}}{U'(W_t)} = \frac{1}{U(W_t)^{\frac{1}{\psi}}} \implies U'(W_t) = \lambda_t C_t^{-\frac{1}{\psi}} U(W_t)^{\frac{1}{\psi}} U'(W_{t+1}) = \lambda_{t+1} C_{t+1}^{-\frac{1}{\psi}} U(W_{t+1})^{\frac{1}{\psi}} \quad (18)$$

Put equation (18) in equation (16) to get

$$\begin{aligned} \lambda_t C_t^{\frac{1}{\psi}} &= \delta (U_{t+1}^*)^{-\frac{1}{\psi}} \left[ E_t \left( U(W_{t+1})^{1-\gamma} \right) \right]^{\frac{1}{1-\gamma}-1} E_t \left( U(W_{t+1})^{-\gamma} \lambda_{t+1} C_{t+1}^{-\frac{1}{\psi}} U(W_{t+1})^{-\frac{1}{\psi}} R_{c,t+1} \right) \\ \lambda_t C_t^{\frac{1}{\psi}} &= \delta (U_{t+1}^*)^{-\frac{1}{\psi}} \frac{U_{t+1}^*}{(U_{t+1}^*)^{1-\gamma}} E_t \left( U(W_{t+1})^{-\gamma} \lambda_{t+1} C_{t+1}^{-\frac{1}{\psi}} U(W_{t+1})^{-\frac{1}{\psi}} R_{c,t+1} \right) \\ \lambda_t C_t^{\frac{1}{\psi}} &= \delta (U_{t+1}^*)^{\gamma-\frac{1}{\psi}} E_t \left( U(W_{t+1})^{-\gamma} \lambda_{t+1} C_{t+1}^{-\frac{1}{\psi}} U(W_{t+1})^{-\frac{1}{\psi}} R_{c,t+1} \right) \\ 1 &= E_t \left( \delta \frac{\lambda_{t+1}}{\lambda_t} \left( \frac{C_{t+1}}{C_t} \right)^{-\frac{1}{\psi}} \frac{(U(W_{t+1}))^{\frac{1}{\psi}-\gamma}}{(U_{t+1}^*)^{\frac{1}{\psi}-\gamma}} \right) \end{aligned}$$

Comparing this with the pricing equation  $1 = E_t(M_{t+1} R_{c,t+1})$ , we get

$$M_{t+1} = \delta \left( \frac{\lambda_{t+1}}{\lambda_t} \right) \left( \frac{C_{t+1}}{C_t} \right)^{-\frac{1}{\psi}} \frac{U(W_{t+1})^{-\frac{1}{\psi}-\gamma}}{(U_{t+1}^*)^{-\frac{1}{\psi}-\gamma}} \quad (19)$$

Guess and verify the policy function

$$U(W_t) = a_t W_t \quad (20)$$

$$C_t = b_t W_t \quad (21)$$

Substitute equations (20) in equation (18),

$$\begin{aligned} a_t &= \lambda C_t^{-\frac{1}{\psi}} (a_t W_t)^{-\frac{1}{\psi}} \\ a_t^{1-\frac{1}{\psi}} &= \lambda_t (b_t W_t)^{-\frac{1}{\psi}} W_t^{\frac{1}{\psi}} \\ a_t^{1-\frac{1}{\psi}} &= \lambda_t b_t^{-\frac{1}{\psi}} \end{aligned} \quad (22)$$

Substituting equation (20) in equation (14), we get

$$a_t W_t = \left[ \lambda_t (b_t W_t)^{1-\frac{1}{\psi}} + \delta \left( [E_t(a_{t+1} W_{t+1})^{1-\gamma}]^{\frac{1}{1-\gamma}} \right)^{1-\frac{1}{\psi}} \right]^{\frac{1}{1-\psi}}$$

$$a_t = \left[ \lambda_t (b_t W_t)^{1-\frac{1}{\psi}} + \delta \left( [E_t(a_{t+1} \frac{W_{t+1}}{W_t})^{1-\gamma}]^{\frac{1}{1-\gamma}} \right)^{1-\frac{1}{\psi}} \right]^{\frac{1}{1-\psi}}$$

From the budget constraint, we have

$$\frac{W_{t+1}}{W_t} = (1 - b_t) R_{a,t+1}$$

Substituting this, we get

$$a_t = \left[ \lambda_t (b_t W_t)^{1-\frac{1}{\psi}} + \delta \left( [E_t(a_{t+1} (1 - b_t) R_{a,t+1})^{1-\gamma}]^{\frac{1}{1-\gamma}} \right)^{1-\frac{1}{\psi}} \right]^{\frac{1}{1-\psi}} \quad (23)$$

Combining equations (22) and (23), we get

$$\lambda_t b_t^{-\frac{1}{\psi}} (1 - b_t) = \delta \left( [E_t(a_{t+1} (1 - b_t) R_{a,t+1})^{1-\gamma}]^{\frac{1}{1-\gamma}} \right)^{1-\frac{1}{\psi}} \quad (24)$$

Plug equations (24) and (22) in equation (19) to get

$$M_{t+1} = \left( \delta \frac{\lambda_{t+1}}{\lambda_t} \right)^{\frac{1-\gamma}{1-\frac{1}{\psi}}} \left( \frac{b_{t+1}}{b_t} (1 - b_t) \right)^{\frac{-(\frac{1}{\psi}-\gamma)}{\psi(1-\frac{1}{\psi})}} \left( \frac{C_{t+1}}{C_t} \right)^{-\frac{1}{\psi}} (R_{a,t+1})^{\frac{1}{\psi}-\gamma}$$

Substituting  $\theta = \frac{1-\gamma}{1-\frac{1}{\psi}}$  and

$$\frac{C_{t+1}}{C_t} \frac{1}{R_{c,t+1}} = \frac{b_{t+1} R_{c,t+1} (W_t - C_t)}{b_t W_t R_{c,t+1}} = \frac{b_{t+1} (1 - b_t)}{b_t}$$

we get the final expression for the stochastic discount factor

$$M_{t+1} = \left( \delta \frac{\lambda_{t+1}}{\lambda_t} \right)^{\theta} \left( \frac{C_{t+1}}{C_t} \right)^{-\frac{\theta}{\psi}} R_{c,t+1}^{\theta-1} \quad (25)$$

Taking log on both sides of equation (25), we get

$$m_{t+1} = \theta \ln \delta + \theta \ln \frac{\lambda_{t+1}}{\lambda_t} - \frac{\theta}{\psi} g_{t+1} + (\theta - 1) r_{a,t+1}$$

$$E_t m_{t+1} = \theta \ln \delta + \theta E_t \ln \frac{\lambda_{t+1}}{\lambda_t} - \frac{\theta}{\psi} E_t g_{t+1} + (\theta - 1) r_{a,t+1}$$

$$m_t - E_t(m_{t+1}) = \theta \left[ \ln \frac{\lambda_{t+1}}{\lambda_t} - E_t \frac{\ln \lambda_{t+1}}{\lambda_t} \right] - \frac{\theta}{\psi} [g_{t+1} - E_t g_{t+1}] + (\theta - 1) (r_{a,t+1} - E_t r_{a,t+1})$$

$$m_t - E_t(m_{t+1}) = \theta [\sigma_{\Lambda} \epsilon_{t+1}^{\Lambda}] - \frac{\theta}{\psi} [\pi_{c,\lambda} \epsilon_{t+1}^{\Lambda} + \sigma_t \eta_{t+1}] + (\theta - 1) [r_{a,t+1} - E_t(r_{a,t+1})] \quad (26)$$

### A.2.1 Measurement Error

Consider the observed consumption growth

$$\tilde{g}_{t+1} = \left( \frac{\tilde{C}_{t+1}}{\tilde{C}_t} \right)^{-\frac{\theta}{\psi}} \quad (27)$$

Consider the true consumption growth

$$g_{t+1} = \left(\frac{C_{t+1}}{C_t}\right)^{-\frac{\theta}{\psi}} \quad (28)$$

where  $\tilde{C}_{t+1} = C_{t+1} + \xi_{t+1}$ ,  $\xi_{t+1} = C_{t+1}^\circ \epsilon_{t+1}$ , and  $\tilde{C}_t = C_t + \xi_t$ ,  $\xi_t = C_t^\circ \epsilon_t$ , with  $\epsilon_t, \epsilon_{t+1} \sim i.i.dN(0, 1)$ .

The Euler Equation is given by

$$\delta^\theta E_t \left[ \left(\frac{\lambda_{t+1}}{\lambda_t}\right)^\theta R_{a,t+1}^{\theta-1} R_{i,t+1} \left(\frac{C_{t+1}}{C_t}\right)^{-\frac{\theta}{\psi}} \right] = 1 \quad (29)$$

The Euler Equation when consumption growth has measurement error is given by

$$\delta^\theta E_t \left[ \left(\frac{\lambda_{t+1}}{\lambda_t}\right)^\theta R_{a,t+1}^{\theta-1} R_{i,t+1} \left(\frac{\tilde{C}_{t+1}}{\tilde{C}_t}\right)^{-\frac{\theta}{\psi}} \right] = p \quad (30)$$

where  $p > 1$ . If there is no measurement error, then the Euler equation is given for  $p = 1$ , which coincides with equation (29).

Expand the observed consumption growth  $(\frac{\tilde{C}_{t+1}}{\tilde{C}_t})^{-\frac{\theta}{\psi}}$  around the point  $C_{t+1}^\circ = E_t[\tilde{C}_{t+1}]$  and  $C_t^\circ = E_{t-1}[\tilde{C}_t]$

$$\begin{aligned} \frac{\tilde{C}_{t+1}}{\tilde{C}_t}^{-\frac{\theta}{\psi}} &\approx \frac{C_{t+1}^\circ}{C_t^\circ} + \frac{1}{\tilde{C}_t^{-\frac{\theta}{\psi}}} \left(-\frac{\theta}{\psi}\right) \bigg|_{(C_{t+1}^\circ, C_t^\circ)} (\tilde{C}_{t+1} - \tilde{C}_{t+1}^\circ) + \tilde{C}_{t+1}^{-\frac{\theta}{\psi}} \left(\frac{\theta}{\psi}\right) \tilde{C}_t^{\frac{\theta}{\psi}-1} \bigg|_{(C_{t+1}^\circ, C_t^\circ)} (\tilde{C}_t - \tilde{C}_t^\circ) \\ &\quad + \frac{1}{2} \frac{1}{\tilde{C}_t^{-\frac{\theta}{\psi}}} \left(\frac{\theta}{\psi}\right) \left(\frac{\theta}{\psi} + 1\right) \tilde{C}_{t+1}^{-\frac{\theta}{\psi}-2} \bigg|_{(C_{t+1}^\circ, C_t^\circ)} (\tilde{C}_{t+1} - \tilde{C}_{t+1}^\circ)^2 \\ &\quad + \frac{1}{2} \tilde{C}_{t+1}^{-\frac{\theta}{\psi}} \left(\frac{\theta}{\psi}\right) \left(\frac{\theta}{\psi} - 1\right) \tilde{C}_{t+1}^{-\frac{\theta}{\psi}-2} \bigg|_{(C_{t+1}^\circ, C_t^\circ)} (\tilde{C}_t - \tilde{C}_t^\circ)^2 \\ &\quad + \left(\frac{\theta}{\psi}\right) \tilde{C}_t^{-\frac{\theta}{\psi}-1} \left(-\frac{\theta}{\psi}\right) \tilde{C}_{t+1}^{-\frac{\theta}{\psi}-1} \bigg|_{(C_{t+1}^\circ, C_t^\circ)} (\tilde{C}_{t+1} - \tilde{C}_{t+1}^\circ)(\tilde{C}_t - \tilde{C}_t^\circ) \\ \frac{\tilde{C}_{t+1}}{\tilde{C}_t}^{-\frac{\theta}{\psi}} &\approx \frac{C_{t+1}^\circ}{C_t^\circ} - \left(\frac{\theta}{\psi}\right) \frac{C_{t+1}^{\circ-\frac{\theta}{\psi}-1}}{C_t^{\circ-\frac{\theta}{\psi}}} (\tilde{C}_{t+1} - \tilde{C}_{t+1}^\circ) + \left(\frac{\theta}{\psi}\right) \frac{C_{t+1}^{\circ-\frac{\theta}{\psi}}}{C_t^\circ} (\tilde{C}_{t+1} - \tilde{C}_{t+1}^{\circ-\frac{\theta}{\psi}+1}) (\tilde{C}_t - \tilde{C}_t^\circ) \\ &\quad + \frac{1}{2} \left(\frac{\theta}{\psi}\right) \left(\frac{\theta}{\psi} + 1\right) \frac{C_{t+1}^{\circ-\frac{\theta}{\psi}-2}}{C_t^{\circ-\frac{\theta}{\psi}}} (\tilde{C}_{t+1} - \tilde{C}_{t+1}^\circ)^2 + \frac{1}{2} \left(\frac{\theta}{\psi}\right) \left(\frac{\theta}{\psi} - 1\right) \frac{C_{t+1}^{\circ-\frac{\theta}{\psi}}}{C_t^{\circ-\frac{\theta}{\psi}+2}} (\tilde{C}_t - \tilde{C}_t^\circ)^2 \\ &\quad - \left(\frac{\theta}{\psi}\right)^2 \frac{C_{t+1}^{\circ-\frac{\theta}{\psi}-1}}{C_t^{\circ-\frac{\theta}{\psi}+1}} (\tilde{C}_{t+1} - \tilde{C}_{t+1}^\circ)(\tilde{C}_t - \tilde{C}_t^\circ) \end{aligned} \quad (31)$$

Similarly, expand the actual consumption growth  $g_{t+1} = (\frac{C_{t+1}}{C_t})^{-\frac{\theta}{\psi}}$  around the point  $C_{t+1}^\circ = E_t[C_{t+1}]$  and  $C_t^\circ = E_{t-1}[C_t]$

$$\begin{aligned} \frac{C_{t+1}}{C_t}^{-\frac{\theta}{\psi}} &\approx \frac{C_{t+1}^\circ}{C_t^\circ} - \left(\frac{\theta}{\psi}\right) \frac{C_{t+1}^{\circ-\frac{\theta}{\psi}-1}}{C_t^{\circ-\frac{\theta}{\psi}}} (C_{t+1} - C_{t+1}^\circ) + \left(\frac{\theta}{\psi}\right) \frac{C_{t+1}^{\circ-\frac{\theta}{\psi}}}{C_t^\circ} (C_{t+1} - C_{t+1}^{\circ-\frac{\theta}{\psi}+1}) (C_t - C_t^\circ) \\ &\quad + \frac{1}{2} \left(\frac{\theta}{\psi}\right) \left(\frac{\theta}{\psi} + 1\right) \frac{C_{t+1}^{\circ-\frac{\theta}{\psi}-2}}{C_t^{\circ-\frac{\theta}{\psi}}} (C_{t+1} - C_{t+1}^\circ)^2 + \frac{1}{2} \left(\frac{\theta}{\psi}\right) \left(\frac{\theta}{\psi} - 1\right) \frac{C_{t+1}^{\circ-\frac{\theta}{\psi}}}{C_t^{\circ-\frac{\theta}{\psi}+2}} (C_t - C_t^\circ)^2 \\ &\quad - \left(\frac{\theta}{\psi}\right)^2 \frac{C_{t+1}^{\circ-\frac{\theta}{\psi}-1}}{C_t^{\circ-\frac{\theta}{\psi}+1}} (C_{t+1} - C_{t+1}^\circ)(C_t - C_t^\circ) \end{aligned} \quad (32)$$

Plug equation (31) in the Euler Equation with measurement error (30).

$$\begin{aligned}
p \approx & \delta^\theta E_t \left[ \left( \frac{\lambda_{t+1}}{\lambda_t} \right)^\theta \left\{ \frac{C_{t+1}^\circ}{C_t^\circ} - \left( \frac{\theta}{\psi} \right) \frac{C_{t+1}^{\circ-\frac{\theta}{\psi}-1}}{C_t^{\circ-\frac{\theta}{\psi}}} (\tilde{C}_{t+1} - \tilde{C}_{t+1}^\circ) + \left( \frac{\theta}{\psi} \right) \frac{C_{t+1}^{\circ-\frac{\theta}{\psi}}}{C_t^\circ} (\tilde{C}_{t+1} - \tilde{C}_{t+1}^{\circ-\frac{\theta}{\psi}+1}) (\tilde{C}_t - \tilde{C}_t^\circ) \right. \right. \\
& + \frac{1}{2} \left( \frac{\theta}{\psi} \right) \left( \frac{\theta}{\psi} + 1 \right) \frac{C_{t+1}^{\circ-\frac{\theta}{\psi}-2}}{C_t^{\circ-\frac{\theta}{\psi}}} (\tilde{C}_{t+1} - \tilde{C}_{t+1}^\circ)^2 + \frac{1}{2} \left( \frac{\theta}{\psi} \right) \left( \frac{\theta}{\psi} - 1 \right) \frac{C_{t+1}^{\circ-\frac{\theta}{\psi}}}{C_t^{\circ-\frac{\theta}{\psi}+2}} (\tilde{C}_t - \tilde{C}_t^\circ)^2 \\
& \left. \left. - \left( \frac{\theta}{\psi} \right)^2 \frac{C_{t+1}^{\circ-\frac{\theta}{\psi}-1}}{C_t^{\circ-\frac{\theta}{\psi}+1}} (\tilde{C}_{t+1} - \tilde{C}_{t+1}^\circ) (\tilde{C}_t - \tilde{C}_t^\circ) \right\} R_{a,t+1}^{\theta-1} R_{i,t+1} \right] \quad (33)
\end{aligned}$$

Similarly, we have

$$\begin{aligned}
1 \approx & \delta^\theta E_t \left[ \left( \frac{\lambda_{t+1}}{\lambda_t} \right)^\theta \left\{ \frac{C_{t+1}^\circ}{C_t^\circ} - \left( \frac{\theta}{\psi} \right) \frac{C_{t+1}^{\circ-\frac{\theta}{\psi}-1}}{C_t^{\circ-\frac{\theta}{\psi}}} (C_{t+1} - C_{t+1}^\circ) + \left( \frac{\theta}{\psi} \right) \frac{C_{t+1}^{\circ-\frac{\theta}{\psi}}}{C_t^\circ} (C_{t+1} - C_{t+1}^{\circ-\frac{\theta}{\psi}+1}) (C_t - \tilde{C}_t^\circ) \right. \right. \\
& + \frac{1}{2} \left( \frac{\theta}{\psi} \right) \left( \frac{\theta}{\psi} + 1 \right) \frac{C_{t+1}^{\circ-\frac{\theta}{\psi}-2}}{C_t^{\circ-\frac{\theta}{\psi}}} (C_{t+1} - C_{t+1}^\circ)^2 + \frac{1}{2} \left( \frac{\theta}{\psi} \right) \left( \frac{\theta}{\psi} - 1 \right) \frac{C_{t+1}^{\circ-\frac{\theta}{\psi}}}{C_t^{\circ-\frac{\theta}{\psi}+2}} (C_t - C_t^\circ)^2 \\
& \left. \left. - \left( \frac{\theta}{\psi} \right)^2 \frac{C_{t+1}^{\circ-\frac{\theta}{\psi}-1}}{C_t^{\circ-\frac{\theta}{\psi}+1}} (C_{t+1} - C_{t+1}^\circ) (C_t - C_t^\circ) \right\} R_{a,t+1}^{\theta-1} R_{i,t+1} \right] \quad (34)
\end{aligned}$$

Subtracting the Euler equations with and without measurement error (33-34) gives us,

$$\begin{aligned}
p - 1 \approx & \delta^\theta E_t \left[ \left( \frac{\lambda_{t+1}}{\lambda_t} \right)^\theta R_{a,t+1}^{\theta-1} R_{i,t+1} \left\{ \frac{\theta}{\psi} \frac{C_{t+1}^{\circ-\frac{\theta}{\psi}-1}}{C_t^{\circ-\frac{\theta}{\psi}}} [C_{t+1} - C_{t+1}^\circ - \tilde{C}_{t+1} + C_{t+1}^\circ] \right. \right. \\
& + \left( \frac{\theta}{\psi} \right) \frac{C_{t+1}^{\circ-\frac{\theta}{\psi}}}{C_t^{\circ-\frac{\theta}{\psi}+1}} [\tilde{C}_{t+1} - C_t^\circ - C_t + C_t^\circ] \\
& + \frac{1}{2} \frac{\theta}{\psi} \left( \frac{\theta}{\psi} + 1 \right) \frac{C_{t+1}^{\circ-\frac{\theta}{\psi}-2}}{C_t^{\circ-\frac{\theta}{\psi}}} [(\tilde{C}_{t+1} - C_{t+1}^\circ) - (C_{t+1} - C_{t+1}^\circ)^2] \\
& + \frac{1}{2} \frac{\theta}{\psi} \left( \frac{\theta}{\psi} - 1 \right) \frac{1}{2} \frac{\theta}{\psi} \left( \frac{\theta}{\psi} + 1 \right) \frac{C_{t+1}^{\circ-\frac{\theta}{\psi}}}{C_t^{\circ-\frac{\theta}{\psi}+2}} [(\tilde{C}_t - C_t^\circ) - (C_t - C_t^\circ)^2] \\
& \left. \left. - \left( \frac{\theta}{\psi} \right)^2 \frac{C_{t+1}^{\circ-\frac{\theta}{\psi}}}{C_t^{\circ-\frac{\theta}{\psi}+1}} [(\tilde{C}_t - C_t^\circ) (\tilde{C}_{t+1} - C_{t+1}^\circ) - (C_t - C_t^\circ) (C_{t+1} - C_{t+1}^\circ)] \right\} \right] \quad (35)
\end{aligned}$$

Substituting the expressions  $\tilde{C}_{t+1} = C_{t+1} + \xi_{t+1}$ ,  $\tilde{C}_t = C_t + \xi_t$  in the above expression (35), we get

$$\begin{aligned}
p - 1 \approx & \delta^\theta \frac{1}{2} \frac{\theta}{\psi} \left( \frac{\theta}{\psi} + 1 \right) \frac{C_{t+1}^{\circ-\frac{\theta}{\psi}-2}}{C_t^{\circ-\frac{\theta}{\psi}}} E_t \left[ \left( \frac{\lambda_{t+1}}{\lambda_t} \right)^\theta R_{a,t+1}^{\theta-1} R_{i,t+1} \right] E_t \left[ 2\xi_{t+1} (C_{t+1} - C_{t+1}^\circ) + C_{t+1}^{\circ 2} \epsilon_{t+1}^2 \right] \\
& + \delta^\theta \frac{1}{2} \frac{\theta}{\psi} \left( \frac{\theta}{\psi} - 1 \right) \frac{C_{t+1}^{\circ-\frac{\theta}{\psi}}}{C_t^{\circ-\frac{\theta}{\psi}+2}} E_t \left[ \left( \frac{\lambda_{t+1}}{\lambda_t} \right)^\theta R_{a,t+1}^{\theta-1} R_{i,t+1} \right] E_t \left[ 2\xi_t (C_t - C_t^\circ) + C_t^{\circ 2} \epsilon_t^2 \right] \\
& - \delta^\theta \left( \frac{\theta}{\psi} \right)^2 \frac{C_{t+1}^{\circ-\frac{\theta}{\psi}}}{C_t^{\circ-\frac{\theta}{\psi}+1}} E_t \left[ \left( \frac{\lambda_{t+1}}{\lambda_t} \right)^\theta R_{a,t+1}^{\theta-1} R_{i,t+1} \right] E_t \left[ \xi_{t+1} (C_t - C_t^\circ) + \xi_t (C_{t+1} - C_{t+1}^\circ) + \xi_t \epsilon_{t+1} \right] \\
p - 1 \approx & \delta^\theta \frac{1}{2} \frac{\theta}{\psi} \left( \frac{\theta}{\psi} + 1 \right) \frac{C_{t+1}^{\circ-\frac{\theta}{\psi}-2}}{C_t^{\circ-\frac{\theta}{\psi}}} E_t \left[ \left( \frac{\lambda_{t+1}}{\lambda_t} \right)^\theta R_{a,t+1}^{\theta-1} R_{i,t+1} \right] \sigma_\epsilon^2 C_{t+1}^{\circ 2} \\
& + \frac{\theta}{\psi} \left( \frac{\theta}{\psi} - 1 \right) \frac{C_{t+1}^{\circ-\frac{\theta}{\psi}}}{C_t^{\circ-\frac{\theta}{\psi}+2}} E_t \left[ \left( \frac{\lambda_{t+1}}{\lambda_t} \right)^\theta R_{a,t+1}^{\theta-1} R_{i,t+1} \right] \sigma_\epsilon^2 C_t^{\circ 2}
\end{aligned}$$



$$\begin{aligned}
&\approx \delta^\theta \frac{1}{2} \left( \frac{\theta}{\psi} \right) \frac{C_{t+1}^{\circ - \frac{\theta}{\psi}}}{C_t^{\circ - \frac{\theta}{\psi}}} E_t \left[ \left( \frac{\lambda_{t+1}}{\lambda_t} \right)^\theta R_{a,t+1}^{\theta-1} R_{i,t+1} \right] \sigma_\epsilon^2 \left( \frac{\theta}{\psi} + 1 + \frac{\theta}{\psi} - 1 \right) \\
p - 1 &\approx \delta^\theta \left( \frac{\theta}{\psi} \right)^2 \sigma_\epsilon^2 E_t \left[ \left( \frac{\lambda_{t+1}}{\lambda_t} \right)^\theta R_{a,t+1}^{\theta-1} R_{i,t+1} \right]
\end{aligned} \tag{36}$$

Therefore, from the expression (36), we get the term  $p$  as

$$\begin{aligned}
p &\approx 1 + \delta^\theta \left( \frac{\theta}{\psi} \right)^2 \sigma_\epsilon^2 E_t \left[ \left( \frac{\lambda_{t+1}}{\lambda_t} \right)^\theta R_{a,t+1}^{\theta-1} R_{i,t+1} \right] \\
&\approx 1 + \Pi \left[ \Omega, R_{i,t+1} \right]
\end{aligned} \tag{37}$$

where  $\Omega$  is a function of the utility parameters  $\Lambda_t, \theta$ , and the aggregate wealth portfolio return  $R_{a,t+1}$ . Thus the modified Euler equation is given as

$$\delta^\theta E_t \left[ \left( \frac{\lambda_{t+1}}{\lambda_t} \right)^\theta R_{i,t+1}^{\theta-1} R_{i,t+1} \left( \frac{\tilde{C}_{t+1}}{\tilde{C}_t} \right)^{-\frac{\theta}{\psi}} \right] \approx 1 + \delta^\theta \left( \frac{\theta}{\psi} \right)^2 \sigma_\epsilon^2 E_t \left[ \left( \frac{\lambda_{t+1}}{\lambda_t} \right)^\theta R_{a,t+1}^{\theta-1} R_{i,t+1} \right] \tag{38}$$

$$\delta^\theta E_t \left[ \left( \frac{\lambda_{t+1}}{\lambda_t} \right)^\theta R_{i,t+1}^{\theta-1} R_{i,t+1} \left( \frac{\tilde{C}_{t+1}}{\tilde{C}_t} \right)^{-\frac{\theta}{\psi}} \right] \approx 1 + \Pi \left[ \Omega, R_{i,t+1} \right] \tag{39}$$

## A.2.2 Representation of Risk free rate

$$E_t [\tilde{M}_{t+1} R_{t+1,f}] = 1 + \Pi[\Omega, R_{f,t+1}]$$

where  $\tilde{M}_{t+1}$  contains the measurement error.

$$\begin{aligned}
1 + \Pi[\Omega, R_{f,t+1}] &= E_t \left[ \exp \left\{ \theta \ln \delta + \theta \ln \frac{\Lambda_{t+1}}{\Lambda_t} - \frac{\theta}{\psi} g_{t+1} + (\theta - 1) r_{a,t+1} + r_{f,t+1} \right\} \right] \\
1 + \Pi[\Omega, R_{f,t+1}] &= \exp \left\{ E_t \left[ \theta \ln \delta + \theta \ln \frac{\Lambda_{t+1}}{\Lambda_t} - \frac{\theta}{\psi} g_{t+1} + (\theta - 1) r_{a,t+1} + r_{f,t+1} \right] \right. \\
&\quad \left. + \frac{1}{2} \text{Var} \left[ \theta \ln \frac{\lambda_{t+1}}{\lambda_t} - \frac{\theta}{\psi} g_{t+1} + (\theta - 1) r_{a,t+1} \right] \right\}
\end{aligned}$$

Taking log on both sides and using the approximation  $\ln(1 + \Pi(\Omega, R_{f,t+1})) \approx \Pi[\Omega, R_{f,t+1}]$ , we get

$$\Pi[\Omega, R_{f,t+1}] \approx \theta \ln \delta - \frac{\theta}{\psi} E_t(g_{t+1}) + \theta E_t \ln \frac{\lambda_{t+1}}{\lambda_t} + (\theta - 1) E_t(r_{a,t+1}) + r_{f,t} + \frac{1}{2} \text{Var}(m_{t+1}) \tag{40}$$

$$\begin{aligned}
E_t(r_a, t+1) &= k_0 + k_1 A_0 + K_1 A_1 E_t(x_{c,t+1}) + k_1 A_2 E_t(x_t^\Lambda) + k_1 A_3 E_t(\sigma_{c,t+1}^2) + k_1 A_4 E_t(\sigma_{x,t+1}^2) \\
&\quad - A_0 - A_1 x_{x,t} - A_2 x_t^\Lambda - A_3 \sigma_{c,t}^2 - A_4 \sigma_{x,t}^2 + E_t(g_{t+1}) \\
E_t(r_a, t+1) &= k_0 + k_1 A_0 + K_1 A_1 \rho x_{c,t} + k_1 A_2 \rho_\Lambda x_t^\Lambda + k_1 A_3 E_t(\sigma_{c,t+1}^2) + k_1 A_4 E_t(\sigma_{x,t+1}^2) \\
&\quad - A_0 - A_1 x_{x,t} - A_2 x_t^\Lambda - A_3 \sigma_{c,t}^2 - A_4 \sigma_{x,t}^2 + x_{c,t} + \mu_c
\end{aligned} \tag{41}$$

$$\begin{aligned}
\text{Var}(m_{t+1}) &= \text{Var} \left[ \theta \ln \delta + \theta \ln \frac{\lambda_{t+1}}{\lambda_t} - \frac{\theta}{\psi} g_{t+1} + (\theta - 1) r_{a,t+1} \right] \\
&= \theta^2 \text{Var} x_{t+1}^\Lambda + \left( \frac{\theta}{\psi} \right)^2 \text{Var}(g_{t+1}) + (\theta - 1)^2 \text{Var}(r_{a,t+1})
\end{aligned} \tag{42}$$

Put equation (41) and equation (42) in equation (40)

$$\begin{aligned}
r_{f,t} = & -\theta \ln \delta + \frac{\theta}{\psi}(\mu_C + x_{c,t}) - \theta \rho_\Lambda \ln \frac{\lambda_t}{\lambda_{t+1}} - (\theta - 1)\{k + 0 + k_1 A_0 + k_1 A_1 \rho x_{c,t} + k_1 A_2 \rho_\Lambda x_t^\Lambda \\
& + k_1 A_3 E_t(\sigma_{c,t+1}^2) + k_1 A_4 E_t(\sigma_{x,t+1}^2) - A_0 - A_1 x_{c,t} - A_2 x_t^\Lambda - A_3 \sigma_{c,t}^2 - A_4 \sigma_{x,t}^2 + x_{c,t} + \mu_c\} \\
& - 0.5\{\theta^2 \text{Var} \sigma_\Lambda^2 + (\frac{\theta}{\psi})^2(\pi_{c,\lambda}^2 + \sigma_{c,t}^2 + (\theta - 1)^2 \text{Var}(r_{a,t+1}))\} - \Pi[\Omega, R_{f,t+1}]
\end{aligned} \tag{43}$$

The risk free rate is thus linear in the state variables  $\mathbf{S}_t = [x_{c,t} \quad x_t^\Lambda \quad \sigma_{x,t}^2 \quad \sigma_{c,t}^2]$ .

The variance of the stochastic discount factor can be decomposed into variances of the state variables.

$$\begin{aligned}
\text{Var}(m_{t+1}) &= \text{Var}\left[\theta x_{t+1}^\Lambda - \frac{\theta}{\psi} g_{t+1} + (\theta - 1)(k_0 + k_1 z_{t+1} - z_t + g_{t+1})\right] \\
&= \text{Var}\left[\theta x_{t+1}^\Lambda - \frac{\theta}{\psi} g_{t+1} + (\theta - 1)z_{t+1} + (\theta - 1)g_{t+1}\right] \\
&= \text{Var}\left[\theta x_{t+1}^\Lambda + \left((\theta - 1) - \frac{\theta}{\psi}\right)g_{t+1} + k_1(\theta - 1)(A_0 + A_1 x_{c,t+1} + A_2 x_{t+1}^\Lambda + A_3 \sigma_{c,t+1}^2) \right. \\
&\quad \left. + A_4 \sigma_{x,t+1}^2\right] \\
&= \text{Var}\left[\theta x_{t+1}^\Lambda + \left((\theta - 1) - \frac{\theta}{\psi}\right)g_{t+1} + k_1(\theta - 1)A_1 x_{c,t+1} + k_1(\theta - 1)A_2 x_{t+1}^\Lambda \right. \\
&\quad \left. + k_1(\theta - 1)A_3 \sigma_{c,t+1}^2 + k_1(\theta - 1)A_4 \sigma_{x,t+1}^2\right] \\
&= [\theta + k_1 A_2(\theta - 1)]^2 \text{Var}(x_{t+1}^\Lambda) + \left[(\theta - 1) - \frac{\theta}{\psi}\right]^2 \text{Var}(g_{t+1}) + [k_1 A_1(\theta - 1)]^2 \text{Var}(x_{c,t+1}) + \\
&\quad [k_1 A_3(\theta - 1)]^2 \text{Var}(\sigma_{c,t+1}^2) + [k_1 A_4(\theta - 1)]^2 \text{Var}(\sigma_{x,t+1}^2)
\end{aligned}$$

This can be written as,

$$\begin{aligned}
\text{Var}(m_{t+1}) &= \lambda_{m,\eta}^2 \text{Var}(g_{t+1}) + \lambda_{m,\lambda}^2 \text{Var}(x_{t+1}^\Lambda) + \lambda_{m,x}^2 \text{Var}(x_{c,t+1}) \\
&\quad + \lambda_{m,\sigma_c}^2 \text{Var}(\sigma_{c,t+1}^2) + \lambda_{m,\sigma_x}^2 \text{Var}(\sigma_{x,t+1}^2)
\end{aligned} \tag{44}$$

where we define the  $\lambda$ 's as,

$$\begin{aligned}
\lambda_{m,\eta} &= (\theta - 1) - \frac{\theta}{\psi} \\
\lambda_{m,\lambda} &= \theta + k_1 A_2(\theta - 1) \\
\lambda_{m,x} &= k_1 A_1(\theta - 1) \\
\lambda_{m,\sigma_c} &= A_3(\theta - 1)k_1 \\
\lambda_{m,\sigma_x} &= A_4(\theta - 1)k_1
\end{aligned}$$

### A.2.3 Deriving Risk Premia

$$\begin{aligned}
E_t[M_{t+1}R_{f,t+1}] &= 1 + \Pi[\Omega, R_{f,t+1}] \\
R_{f,t+1} &= \frac{1 + \Pi[\Omega, R_{f,t+1}]}{E_t[M_{t+1}]} \\
\ln R_{f,t+1} &= \ln[1 + \Pi[\Omega, R_{f,t+1}]] - \ln E_t M_{t+1} \\
r_{f,t+1} &= \Pi[\Omega, R_{f,t+1}] - \ln E_t e^{\ln M_{t+1}} \\
r_{f,t+1} &= \Pi[\Omega, R_{f,t+1}] - \ln e^{E_t[m_{t+1}] + \frac{1}{2} \text{Var}(m_{t+1})} \\
r_{f,t+1} &= \Pi[\Omega, R_{f,t+1}] - E_t(m_{t+1}) - \frac{1}{2} \text{Var}(m_{t+1}) \\
E_t[M_{t+1}R_{a,t+1}] &= 1 + \Pi[\Omega, R_{f,t+1}] \\
e^{E_t(m_{t+1}) + E_t r_{a,t+1} + \frac{1}{2} \text{Var}(m_{t+1} + r_{a,t+1})} &= 1 + \Pi[\Omega, R_{f,t+1}]
\end{aligned} \tag{45}$$

Taking log on both sides and using the approximation  $\ln(1 + \Pi[\Omega, R_{f,t+1}]) \approx \Pi[\Omega, R_{f,t+1}]$ , we get

$$\begin{aligned}
\Pi[\Omega, R_{f,t+1}] &= E_t(m_{t+1}) + E_t(r_{a,t+1}) + \frac{1}{2} \text{Var}(m_{t+1} + r_{a,t+1}) \\
E_t(m_{t+1}) &= -E_t(r_{a,t+1}) - \frac{1}{2} [\text{Var}(m_{t+1}) + \text{Var}(r_{a,t+1}) + 2\text{Cov}(m_{t+1}, r_{a,t+1})] \\
&\quad + \Pi[\Omega, R_{f,t+1}] \\
E_t(m_{t+1}) + \frac{1}{2} \text{Var}(r_{m,t+1}) &= -E_t(r_{a,t+1}) - \frac{1}{2} [\text{Var}(r_{a,t+1}) + 2\text{Cov}(m_{t+1}, r_{a,t+1})]
\end{aligned} \tag{46}$$

From equation (45), we get

$$\begin{aligned}
E_t(r_{a,t+1} - r_f) &= \Pi[\Omega, R_{f,t+1}] - \frac{1}{2} [\text{Var}(r_{a,t+1}) + 2\text{Cov}(m_{t+1}, r_{a,t+1})] \\
E_t(r_{a,t+1} - r_f) + 0.5 \text{Var}(r_{a,t+1}) &= \Pi[\Omega, R_{f,t+1}] + \text{Cov}[(m_{t+1} - E_t(m_{t+1})), (r_{a,t+1} - E_t(r_{a,t+1}))]
\end{aligned} \tag{47}$$

Asset returns follows the process,

$$\begin{aligned}
r_{a,t+1} &= k_0 + k_1 z_{t+1} - z_t + g_{t+1} \\
&= k_0 + k_1 [A_0 + A_1 x_{c,t+1} + A_2 x_{t+1}^\Lambda + A_3 \sigma_{c,t+1}^2 + A_4 \sigma_{x,t+1}^2] \\
&\quad - A_0 - A_1 x_{c,t} - A_2 x_t^\Lambda - A_3 \sigma_{c,t}^2 - A_4 \sigma_{x,t}^2 + g_{t+1}
\end{aligned} \tag{48}$$

$$\begin{aligned}
E_t(r_{a,t+1}) &= k_0 + k_1 A_0 + K_1 A_1 E_t(x_{c,t+1}) + k_1 A_2 E_t(x_{t+1}^\Lambda) + k_1 A_3 E_t(\sigma_{c,t+1}^2) + k_1 A_4 E_t(\sigma_{x,t+1}^2) \\
&\quad - A_0 - A_1 x_{c,t} - A_2 x_t^\Lambda - A_3 \sigma_{c,t}^2 - A_4 \sigma_{x,t}^2 + E_t(g_{t+1})
\end{aligned} \tag{49}$$

$$\begin{aligned}
r_{a,t+1} - E_t(r_{a,t+1}) &= k_1 A_1 [x_{c,t+1} - E_t(x_{c,t+1})] + k_1 A_2 [x_{t+1}^\Lambda - E_t(x_{t+1}^\Lambda)] + k_1 A_3 [\sigma_{c,t+1}^2 - E_t(\sigma_{c,t+1}^2)] \\
&\quad + k_1 A_4 [\sigma_{x,t+1}^2 - E_t(\sigma_{x,t+1}^2)] + g_{t+1} - E_t(g_{t+1}) \\
&= k_1 A_1 \varphi_e \sigma_{x,t} e_{t+1} + k_1 A - 2\sigma_\Lambda \epsilon_{t+1}^\Lambda + k_1 A_3 \omega_{c,t+1} + k_1 A_4 \omega_{x,t+1} + \sigma_{c,t} \eta_{t+1} + \pi_{c,\lambda} \epsilon_{t+1}^\Lambda \\
&= k_1 A_1 \varphi_e \sigma_{x,t} e_{t+1} + (k_1 A_2 \sigma_\Lambda + \pi_{c,\lambda}) \epsilon_{t+1}^\Lambda + k_1 A_3 \omega_{c,t+1} + k_1 A_4 \omega_{x,t+1} + \sigma_{c,t} \eta_{t+1}
\end{aligned} \tag{50}$$

note that  $\omega_{c,t+1} = \sigma_{c,t+1}^2 - E_t(\sigma_{c,t+1}^2)$

Recall equation (26)

$$m_t - E_t(m_{t+1}) = \theta[\sigma_\Lambda \epsilon_{t+1}^\Lambda] - \frac{\theta}{\psi}[\pi_{c,\lambda} \epsilon_{t+1}^\Lambda + \sigma_{c,t} \eta_{t+1}] + (\theta - 1)[r_{a,t+1} - E_t(r_{a,t+1})]$$

Substitute equation (50) in equation (26),

$$\begin{aligned} m_t - E_t(m_{t+1}) &= \theta \sigma_\Lambda \epsilon_{t+1}^\Lambda - \frac{\theta}{\psi}[\pi_{c,\lambda} \epsilon_{t+1}^\Lambda + \sigma_{c,t} \eta_{t+1}] + (\theta - 1)[k_1 A_1 \varphi_e \sigma_{x,t} e_{t+1} + (k_1 A_2 \sigma_\Lambda + \pi_{c,\lambda}) \epsilon_{t+1}^\Lambda \\ &\quad + k_1 A_3 \omega_{c,t+1} + k_1 A_4 \omega_{x,t+1} + \sigma_{c,t} \eta_{t+1}] \\ &= [\theta \sigma_\Lambda - \frac{\theta}{\psi} \pi_{c,\lambda} + (\theta - 1)(k_1 A_2 \sigma_\Lambda + \pi_{c,\lambda})] \epsilon_{t+1}^\Lambda + [(\theta - 1) - \frac{\theta}{\psi} \sigma_{c,t} \eta_{t+1} \\ &\quad + k_1 A_1 \varphi - e(\theta - 1) \sigma_{x,t} e_{t+1} + k_1 A_3 (\theta - 1) \omega_{c,t+1} + k_1 A_4 (\theta - 1) \omega_{x,t+1}] \\ &= \left[ \sigma_\Lambda (\theta + (\theta - 1) k_1 A_2) + \pi_{c,\lambda} [(\theta - 1) \frac{\theta}{\psi}] \right] \epsilon_{t+1}^\Lambda + [(\theta - 1) - \frac{\theta}{\psi} \sigma_{c,t} \eta_{t+1} \\ &\quad + k_1 A_1 (\theta - 1) \varphi_e \sigma_{x,t} e_{t+1} + k_1 A_3 (\theta - 1) \omega_{c,t+1} + k_1 A_4 (\theta - 1) \omega_{x,t+1}] \end{aligned}$$

Comparing with the  $\lambda$ 's from equation (44), we can write the innovation to the stochastic discount factor in the above equation as as,

$$\begin{aligned} m_t - E_t(m_{t+1}) &= [\sigma_\Lambda \lambda_{m,\lambda} + \pi_{c,\lambda} \lambda_{m,\eta}] \epsilon_{t+1}^\Lambda + \lambda_{m,\eta} \sigma_{c,t} \eta_{t+1} + \lambda_{m,x} \varphi_e \sigma_{x,t} e_{t+1} \\ &\quad + \lambda_{m,\sigma_c} \omega_{c,t+1} + \lambda_{m,\sigma_x} \omega_{x,t+1} \end{aligned}$$

We notice that the proportion of  $\epsilon_{t+1}^\Lambda$  (noise in the time preference variable) that gets attributed to innovation in stochastic discount factor is the weighted average of  $\lambda_{m,\eta}$  and  $\lambda_{m,\lambda}$  with weights  $\pi_{c,\lambda}$ , and  $\sigma_\Lambda$  respectively. This is because  $\epsilon_{t+1}^\Lambda$  enters both the time preference dynamics as well as the consumption growth process. We can define  $\lambda_{m,x^\Lambda} = [\lambda_{m,\lambda} + \frac{\pi_{c,\lambda}}{\sigma_\Lambda} \lambda_{m,\eta}]$  and rewrite the innovation to stochastic discount factor as

$$m_t - E_t(m_{t+1}) = \lambda_{m,x^\Lambda} \sigma_\Lambda \epsilon_{t+1}^\Lambda + \lambda_{m,\eta} \sigma_{c,t} \eta_{t+1} + \lambda_{m,x} \varphi_e \sigma_{x,t} e_{t+1} + \lambda_{m,\sigma_c} \omega_{c,t+1} + \lambda_{m,\sigma_x} \omega_{x,t} \quad (51)$$

Recall equation (47)

$$E_t(r_{a,t+1} - r_f) + 0.5 \text{Var}(r_{a,t+1}) = \Pi[\Omega, R_{f,t+1}] + \text{Cov}[(m_{t+1} - E_t(m_{t+1})), r_{a,t+1} - E_t(r_{a,t+1})]$$

$$\begin{aligned} E_t(r_{a,t+1} - r_f) + 0.5 \text{Var}(r_{a,t+1}) &= -\text{Cov} \left\{ \left( \lambda_{m,x^\Lambda} \sigma_\Lambda \epsilon_{t+1}^\Lambda + \lambda_{m,\eta} \sigma_{c,t} \eta_{t+1} + \lambda_{m,x} \varphi_e \sigma_{x,t} e_{t+1} + \lambda_{m,\sigma_c} \omega_{c,t+1} \right. \right. \\ &\quad \left. \left. + \lambda_{m,\sigma_x} \omega_{x,t+1} \right), r_{a,t+1} - E_t(r_{a,t+1}) \right\} + \Pi[\Omega, R_{a,t+1}] - \Pi[\Omega, R_{f,t+1}] \end{aligned} \quad (52)$$

Let  $\tilde{r}_{a,t+1} = r_{a,t+1} - E_t(r_{a,t+1})$

$$\begin{aligned}
E_t(r_{a,t+1} - r_f) + 0.5Var(r_{a,t+1}) = & - \left\{ \lambda_{m,x^\Lambda} Cov(\sigma_\Lambda \epsilon_{t+1}^\Lambda, \tilde{r}_{a,t+1}) + \lambda_{m,\eta} Cov(\sigma_{c,t} \eta_{t+1}, \tilde{r}_{a,t+1}) \right. \\
& + \lambda_{m,x} Cov(\varphi_e \sigma_{x,t} e_{t+1}, \tilde{r}_{a,t+1}) + \lambda_{m,\sigma_c} Cov(\omega_{c,t+1}, \tilde{r}_{a,t+1}) \\
& \left. + \lambda_{m,\sigma_x} Cov(\omega_{x,t+1}, \tilde{r}_{a,t+1}) \right\} + \Pi[\Omega, R_{a,t+1}] - \Pi[\Omega, R_{f,t+1}]
\end{aligned} \tag{53}$$

This forms the basic testing equation for cross sectional asset returns in the case of omitted factors. Note that the bias induced from the measurement error enters this testing equation, which may bias the risk premia estimates if not accounted for. The corresponding first stage estimation equation is the given by

$$r_{a,t+1} = cons_a + \beta_{a,\Lambda} \epsilon_{t+1}^\Lambda + \beta_{a,\eta} \sigma_t \eta_{t+1} + \beta_{a,x} \varphi_e \sigma_t e_{t+1} + \beta_{a,\omega_c} \omega_{c,t+1} + \beta_{a,\omega_x} \omega_{x,t+1} \tag{54}$$

where the  $\beta$ 's are the covariance of the corresponding factors with the innovation in asset returns.

### A.3 Empirics

### A.3.1 News Index



Figure 7: Word Cloud: Words the occur at least 150 times in all articles combined. Bigger size implies higher frequency.

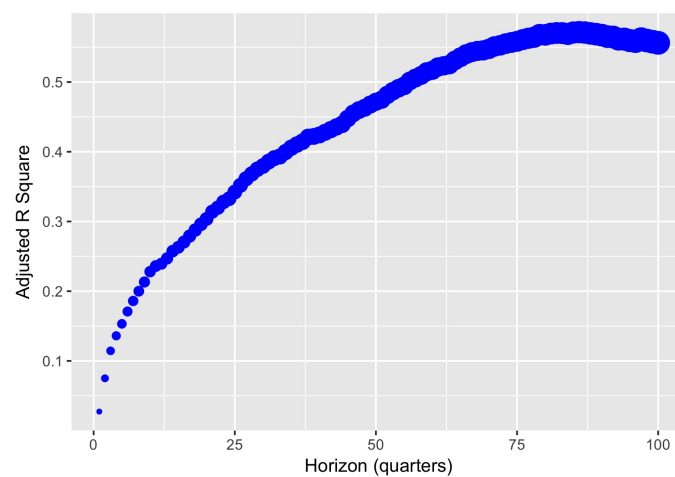


Figure 8: Consumption growth predictability

### A.3.2 Machine Learning model performance

Table 3: Model Performance comparison

	Test Sucess Rate	Training Sucess Rate
Ridge	77%	-
Lasso	78%	-
Classification Tree	72%	75%
Bagging	77%	79%
Random Forest	81%	82%
Support Vector Machines	70%	71%

### A.3.3 Predictability Results

The predictability regression equation is given by

$$\sum_{i=1}^K g_{t+i} = \beta_0 + \beta_1 \text{NewsIndex}_t + \epsilon_{t+K}, \quad K = 1, 2, \dots, 100$$

where  $K$  indicates the months and  $g_{t+i}$  represents the consumption growth between periods  $t+i$  and  $t+i-1$ .

Table 4: Simple Regression

Time Horizon	1	20	40	60	80	100
Slope: Index	-0.0005***	-0.011***	-0.02***	-0.03***	-0.04***	-0.06***
Adj R2	3%	30%	42%	52%	57%	56%

\*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.10$

The multivariate regression includes the price dividend ratio, real labor income growth, and realized industrial production variance as controls. The equation is given by

$$\sum_{i=1}^K g_{t+i} = \beta_0 + \beta_1 \text{NewsIndex}_t + \text{Controls}_t + \epsilon_{t+K}, \quad K = 1, 2, \dots, 100$$

Table 5: Multiple Regression

Time Horizon	1	20	40	60	80
Slope: Index	-0.0007***	-0.013***	-0.02***	-0.03***	-0.04***
Slope: Dividend Price Ratio	0.001***	0.0097***	0.011***	0.013***	0.009*
Slope: Labor Income	0.08***	0.18*	0.20	0.15	0.23
Slope: Industrial Production Variance	-0.03	-0.09	0.87	4.85	7.3
Adj R2	7%	32%	43%	53%	58%

\*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.10$



#### A.3.4 Fitting consumption growth

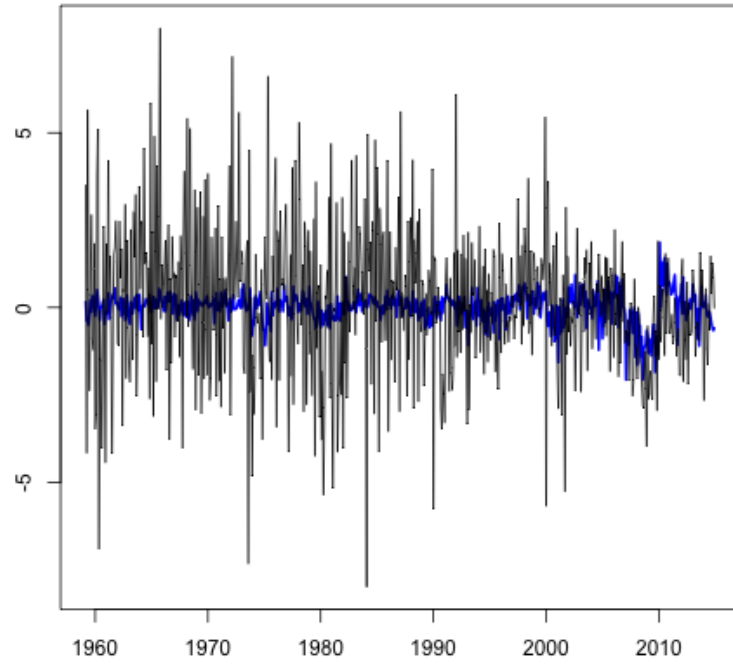


Figure 9: Consumption growth against persistent components

### A.3.5 Measurement Error

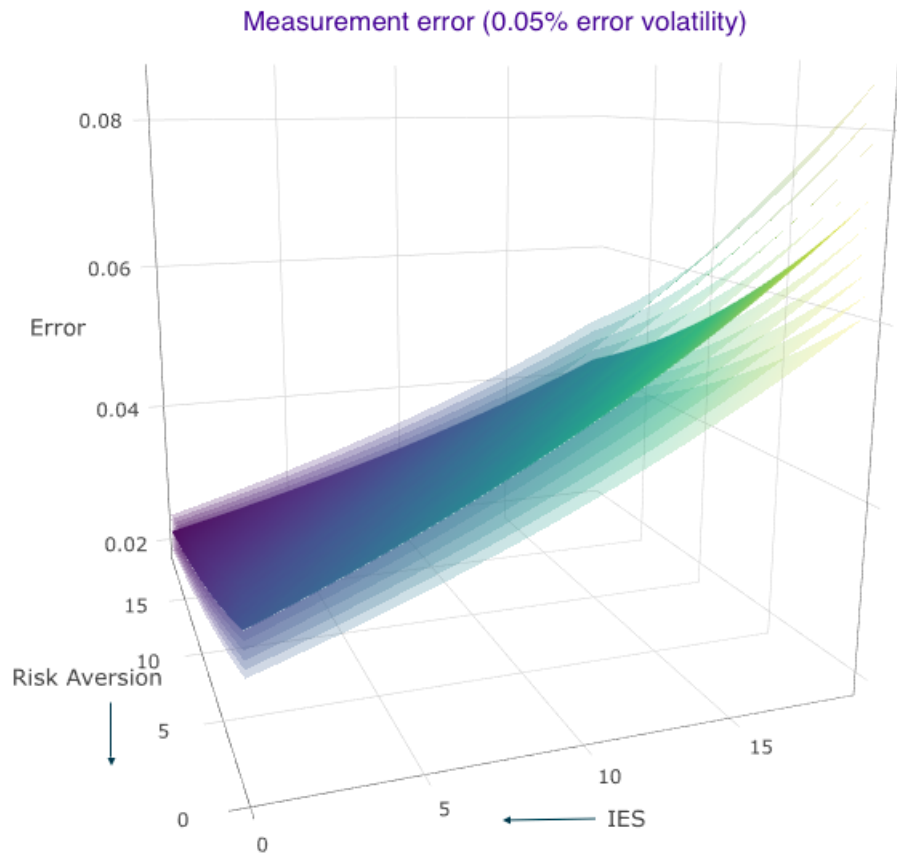


Figure 10: Measurement error with variance of 0.005%. Risk aversion is increasing from top to bottom, IES is increasing from right to left.

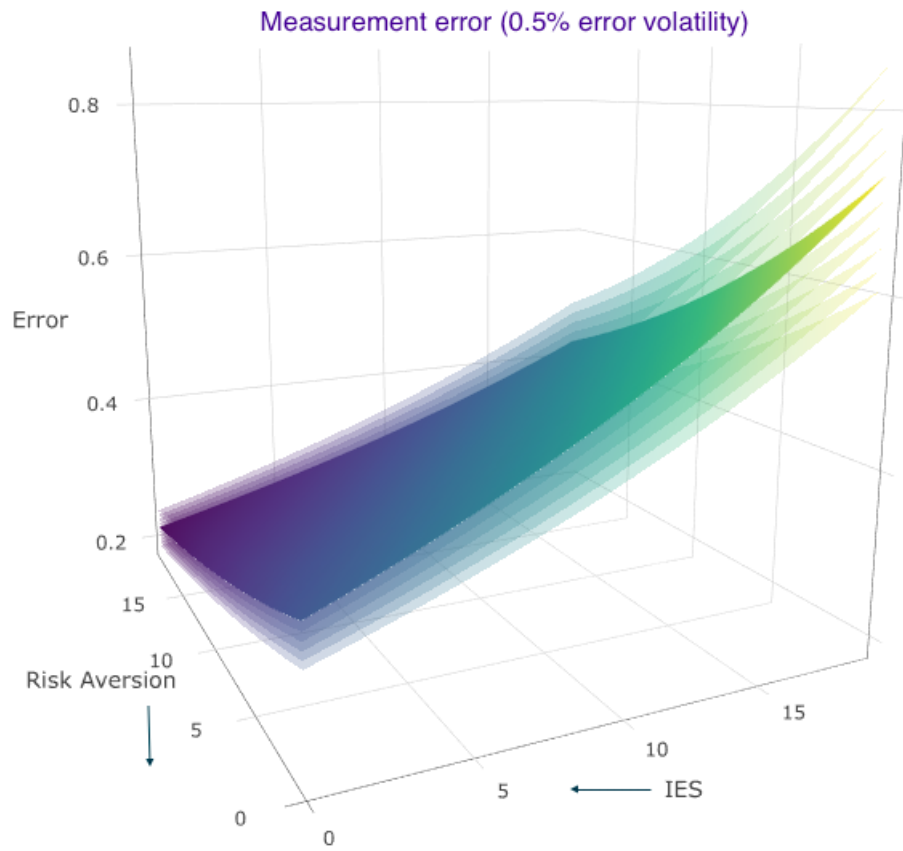


Figure 11: Measurement error with variance of 0.05%. Risk aversion is increasing from top to bottom, IES is increasing from right to left.

### A.3.6 Three pass Estimation: Second stage Results

	Model 1
(Intercept)	1.53*** (0.29)
PC1	0.13 (0.07)
PC2	−0.03** (0.01)
PC3	0.01* (0.01)
PC4	−0.04*** (0.01)
PC5	−0.02** (0.00)
PC6	0.02** (0.01)
PC7	−0.01 (0.01)
R <sup>2</sup>	0.79
Adj. R <sup>2</sup>	0.75
Num. obs.	45
RMSE	0.12

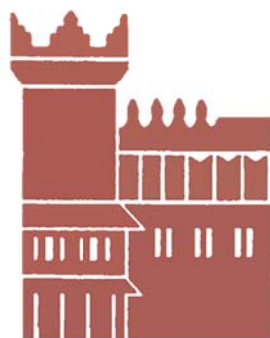
\*\*\*  $p < 0.001$ , \*\*  $p < 0.01$ , \*  $p < 0.05$

Table 6: Stage 2 results

Step 2 (CSR): Obtain the risk premia of estimated factor scores by running a cross sectional regression of average asset returns onto estimated factor loadings  $\hat{\beta}$ . The risk premia of the estimated factor returns is given by

$$\tilde{\Gamma} := (\tilde{\gamma}_0, \tilde{\gamma}^T)^T = \left( (\iota_n : \hat{\beta})^T (\iota_n : \beta) \right)^{-1} (\iota_n : \hat{\beta})^T \bar{r}$$

The cross sectional regression of asset returns on the estimated factor scores provide an R-squared of 75%. Such high R-squared is not surprising since the factor scores are obtained from the statistical decomposition of the asset returns in the first place.



Alma Mater Studiorum - Università di Bologna  
DEPARTMENT OF ECONOMICS

Strada Maggiore 45  
40125 Bologna - Italy  
Tel. +39 051 2092604  
Fax +39 051 2092664  
<http://www.dse.unibo.it>